

3)

Journal of
**Geophysical
Research**

VOLUME 66

OCTOBER 1961

NUMBER 10

PUBLISHED BY

THE AMERICAN GEOPHYSICAL UNION

Journal of Geophysical Research

An International Scientific Publication

OFFICERS OF THE UNION

THOMAS F. MALONE, *President*
GEORGE P. WOOLLARD, *Vice President*
A. NELSON SAYRE, *General Secretary*
WALDO E. SMITH, *Executive Secretary*

OFFICERS OF THE SECTIONS

Geodesy

FLOYD W. HOUGH, *President*
CHARLES A. WHITTEN, *Vice President*
BUFORD K. MEADE, *Secretary*

Seismology

JAMES A. PEOPLES, JR., *President*
JACK E. OLIVER, *Vice President*
BENJAMIN F. HOWELL, JR., *Secretary*

Meteorology

MORRIS NEIBURGER, *President*
HENRY G. HOUGHTON, *Vice President*
WOODROW C. JACOBS, *Secretary*

Geomagnetism and Aeronomy

C. T. ELVEY, *President*
E. H. VESTINE, *Vice President*
J. HUGH NELSON, *Secretary*

Oceanography

DONALD W. PRITCHARD, *President*
ROBERT S. ARTHUR, *Vice President*
ARTHUR E. MAXWELL, *Secretary*

Volcanology, Geochemistry, and Petrology

HATTEN S. YODER, JR., *President*
EDWARD D. GOLDBERG, *Vice President*
DAVID R. WONES, *Secretary*

Hydrology

WILLIAM C. ACKERMANN, *President*
DAVID K. TODD, *Vice President*
RALPH N. WILSON, *Secretary*

Tectonophysics

LOUIS B. SLICHTER, *President*
DAVID T. GRIGGS, *Vice President*
IRIS BORG, *Secretary*

BOARD OF EDITORS

Editors: PHILIP H. ABELSON and J. A. PEOPLES,

ASSOCIATE EDITORS

1959-1961

HENRI BADER	T. NAGATA
K. E. BULLEN	FRANK PRESS
CONRAD P. MOOK	A. NELSON SAYRE
WALTER H. MUNK	MERLE A. TUVE
JAMES A. VAN ALLEN	

1960-1962

JULIUS BARTELS	L. A. MANNING
V. V. BELOUSSOV	TOR J. NORDENSON
E. G. BOWEN	E. N. PARKER
JOHN E. CHAPPELEAR	GEORGE P. RIGSBY
G. D. GARLAND	WALTER O. ROBERTS
GORDON J. F. MACDONALD	C. N. TOUART
JAMES R. WAIT	

1961-1963

FRANKLIN I. BADGLEY	ROBERT O. REID
HENRY G. BOOKER	BRUNO ROSSI
JOSEPH W. CHAMBERLAIN	GEORGE H. SUTTON
HERBERT FRIEDMAN	DAVID K. TODD
MARK F. MEIER	VICTOR VACQUIER
ARTHUR H. WAYNICK	

The Editors of the *Journal of Geophysical Research* welcome original scientific contributions on the physics of the earth and its environment.

Manuscripts should be submitted in triplicate. J. A. Peoples, Jr., Department of Geology, University of Kansas, Lawrence, Kansas. Authors' institutions if in the United States or Canada, are requested to pay a publication charge of \$25 per page, which, honored, entitles them to 100 free reprints.

Subscriptions to the *Journal of Geophysical Research* and *Transactions, AGU*, are included in membership dues.

Nonmember subscriptions, *Journal of Geophysical Research*, \$30 for back volume of 1959, \$42 for back volume of 1960, \$6 for this issue; \$20 for the calendar year 1961.

Nonmember subscriptions, *Transactions, AGU*, \$4 per calendar year, \$1.25 per copy.

Subscriptions, renewals, and orders for back numbers should be addressed to American Geophysical Union, 1515 Massachusetts Ave., Northwest, Washington, D. C. Suggestions to authors are available on request.

Advertising Representative: Howland and Howland Inc., 230 Park Ave., New York 17, N. Y.

Since January 1959 (Vol. 64, No. 1) the *Journal of Geophysical Research* has been published monthly by the American Geophysical Union, the U. S. National Committee of the International Union of Geodesy and Geophysics, organized under the National Academy of Sciences-National Research Council as the U. S. national adhering body. Publication of this journal is supported by the National Science Foundation and the Carnegie Institution of Washington. The new monthly combines the type of scientific material formerly published in the bi-monthly *Transactions, American Geophysical Union*, and the quarterly *Journal of Geophysical Research*. The *Transactions, American Geophysical Union*, will continue as a quarterly publication for Union business and items of interest to members of the Union.

Published monthly by the American Geophysical Union from 1407 Sherwood Avenue, Richmond, Virginia. Second class postage paid at Richmond, Virginia.

only LaCoste & Romberg
Geodetic Gravity Meter

gives you **thermal controlled accuracy in a 7-pound meter**

- world survey without resetting
- never requires recalibrating
- less than 0.5 mg/l drift per month
- no "sets" or "tares" under normal operation

This new miniaturized Geodetic Gravity Meter retains all the accuracy and dependability of the standard model introduced by LaCoste & Romberg in 1956, yet it weighs only 7 pounds. (Complete with battery and luggage-type carrying case, it weighs less than 17 pounds). With a world-wide range of over 6,000 mg/l., this instrument has a repeatability of 0.01 mg/l. Actual field tests over the complete gravity range have shown an accuracy better than 0.04 mg/l.

Exceptionally high sensitivity of the LaCoste & Romberg meter is attained by a zero length spring suspension (U. S. Patent No. 2,293,437). Calibration is stabilized by means of patented

lever systems that act on the main spring rather than on weak measuring springs. And by thermostating, drift is normally reduced to less than 0.5 milligal per month.

Rugged and dependable, the LaCoste & Romberg Geodetic Meter requires practically no maintenance in the field. Its gravity responsive system is completely suspended by springs and will therefore withstand any shock that will not damage the housing supporting it. It is specifically designed to provide a light-weight meter with higher accuracy and lower drift than can be attained in any other geodetic gravity meter. For complete information, write for *Miniature Geodetic Gravity Meter Bulletin*.



LaCoste & Romberg

6606 NORTH LAMAR

AUSTIN, TEXAS

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers

SPRENGNETHER LONG PERIOD HORIZONTAL SEISMOMETER

GENERAL SPECIFICATIONS:

- **Period Range:** 6 to 70 seconds.
- **Magnification:** Up to 15,000, depending on operating period.
- **Damping:** Electromagnetic.
- **Transducer-moving coils in circular magnetic gaps.**
- **Coils approximately 500 ohms. or to your specifications.**
- **Pendulum steady mass weight 22 lbs.**

PHYSICAL SPECIFICATIONS:

Length.....	25"
Width.....	25"
Height.....	22"
Net Weight.....	90 lbs. approx.
Shipping Weight.....	155 lbs. approx.

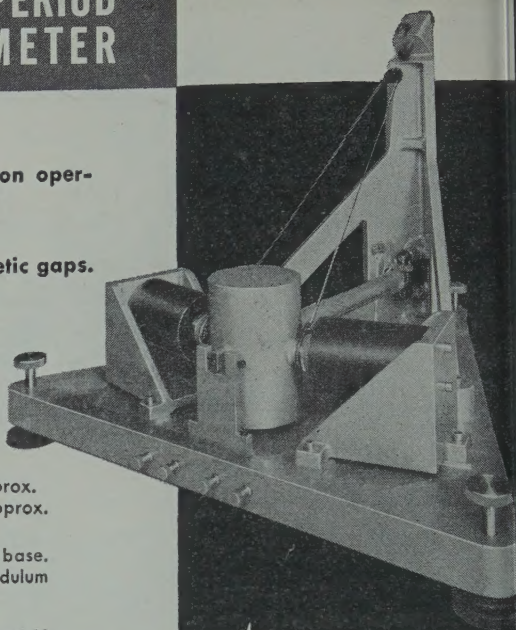
COVER:

Light weight aluminum with rubber gasket seal to base.
Provided with viewing window to observe pendulum centering.

*To compliment this instrument,
a long period vertical seismometer is also available.*

Internationally Known Mfrs. of Seismological, Geophysical Instruments.

W. F. SPRENGNETHER INSTRUMENT CO., INC.
4567 SWAN AVENUE. ST. LOUIS 10, MO.



WRITE FOR DETAILED
INFORMATION ON THESE
INSTRUMENTS.

GEODESY AND CARTOGRAPHY, 1959 and 1960

(Geodezia i Kartografiya)

The leading monthly journal of Geodesy and Cartography in the USSR is being translated and published in an English edition, *for the years 1959 and 1960*, by the American Geophysical Union, aided by a grant from the National Science Foundation.

Subscription price, \$20.00 for each volume of 12 numbers

Send subscription requests to

AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W.

Washington 5, D. C., U.S.A.

Other publications of the AGU include

- JOURNAL OF GEOPHYSICAL RESEARCH (monthly), 4800 pages anticipated for 1961, subscription \$20.00
- TRANSACTIONS (quarterly), \$4.00 per calendar year
- GEOPHYSICAL MONOGRAPH SERIES (occasional volumes), No. 5 (issued in 1960) *Physics of Precipitation* (proceedings of the Cloud Physics Conference, Woods Hole, Massachusetts, June 3-5, 1959), 435 pp., \$12.50
- IZVESTIYA, Academy of Sciences, USSR, Geophysics series, English edition (monthly), \$25.00 per calendar year, available for 1957, 1958, 1959, 1960, and 1961.

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers

IONOSPHERIC PROPAGATION and HF COMMUNICATIONS



If you have the background, the imagination and the desire to contribute to important programs in these fields, you are invited to join a carefully selected team of outstanding scientists and engineers now contributing significantly to current knowledge through advanced research.

These programs are being conducted at our **ELECTRO-PHYSICS LABORATORIES** in the **SUBURBAN WASHINGTON, D. C.** area, ideally located from the viewpoint of advanced study which may be conducted at one of several nearby universities; for readily available housing in pleasant residential neighborhoods; and for the general amenities of living offered by this important Metropolitan center. All qualified applicants will receive consideration for employment without regard to race, creed, color or national origin.

OUR PRESENT NEEDS ARE FOR:

SENIOR IONOSPHERIC

PHYSICISTS Ph.D. preferred, with several years' experience in the study of Ionospheric phenomena. Should be familiar with present knowledge of upper atmosphere physics and possess an understanding of current programs using rockets and satellites for studies in F-region and beyond. Qualified individuals with supervisory abilities will have an exceptional opportunity to assume project leadership duties on HF projects already under way involving F-layer propagation studies, backed by a substantial experimental program.

SENIOR DEVELOPMENT

PHYSICISTS Advanced degree in Physics or E.E. preferred. Must be familiar with latest techniques in the design of advanced HF receivers and transmitters and possess working knowledge of modern HF networks employing ferrites and metallic tape cores. Strong theoretical background in modern linear circuit theory desired. Will carry out laboratory development and implementation of new HF communications systems.

SENIOR ELECTRONIC

ENGINEERS Advanced degree in E.E. preferred. Must be familiar with conventional pulse circuit designs and applications. Technical background should include substantial experience in data process and data recovery systems using

both analog and digital techniques. Knowledge of principles and application of modern information theory including correlation techniques helpful. Will be responsible for the design of sub-systems.

JUNIOR ELECTRONIC

ENGINEERS To assist Senior Engineers and Scientists in the development of HF communications and data process equipment. Should have formal electronics schooling and 2 years' experience in circuit design checkout or analysis of HF communications, Radar Pulse, Analog/Digital or Data Recovery equipment. Construction of prototypes of new and interesting equipment and design of individual components of communications and data processing systems will comprise the major efforts of selected applicants.

FIELD STATION

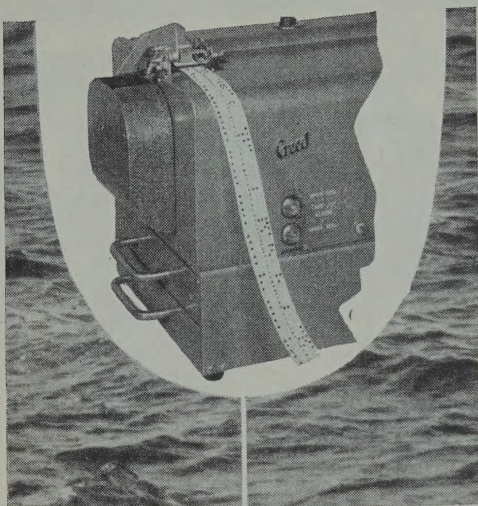
ENGINEERS B.S.E.E. or equivalent, consisting of combined civilian or military technical school, with work experience. Presently employed as a field engineer or project engineer with a valid 1st or 2nd Class FCC license and a good command of some of the following: Radar, preferably high power; HF long-distance communications systems; Tropospheric or Ionospheric scatter systems. Must be willing to accept assignments in areas where dependents are not permitted for periods of up to one year. Differential paid for overseas assignments.

For a prompt reply to your inquiry, please forward resume in confidence to: Dept. C-1, W. T. WHELAN, Director of Research & Development

ACF ELECTRONICS
DIVISION

ACF INDUSTRIES

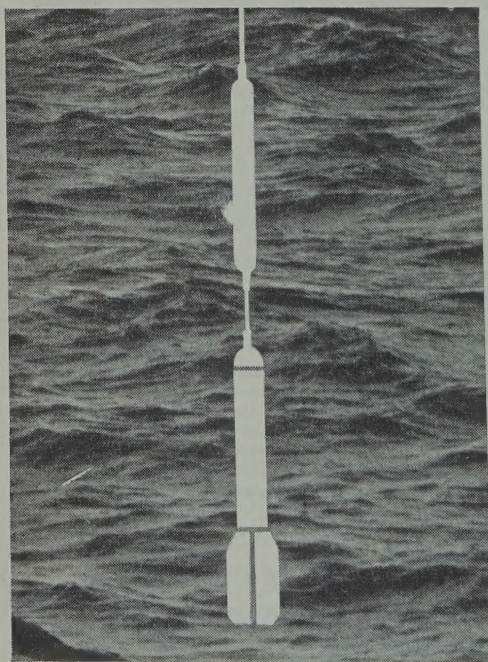
HYATTSVILLE, MARYLAND



SHIPBORNE PROTON SPIN MAGNETOMETER

for Geo-Magnetic Survey Work

A Resolution of 0.5 in 50,000 gammas
Analogue and Digital Recording Facilities
Six-meter Decimal Display Count



Bruce Peebles & Co. Ltd
EDINBURGH, SCOTLAND

ARE YOU A MEMBER of AGU?

If you are a reader or user of either of the publications noted below, you should look into this matter.

Members regularly receive
the monthly

**Journal of
Geophysical Research**

the quarterly

**Transactions
American Geophysical Union**

Members are entitled to special discounts on other publications of AGU such as

- **Geophysical Monograph Series** (Nos. 1 to 5 available)
 - **Izvestiya of the Academy of Sciences, USSR; Geophysics Series** for the years 1957, 1958, 1959 and 1960
 - **Geodesy and Cartography, USSR** 1959
-
- Annual meeting in Washington in late April or early May each year
 - Regional meeting in the Pacific Northwest in autumn
 - Regional meeting in the Pacific Southwest in winter
 - National Western Meeting, December 1961

For application forms and other information regarding membership meetings, and other matters, write to

AMERICAN GEOPHYSICAL UNION
1515 Massachusetts Ave., N.W.
Washington 5, D. C.

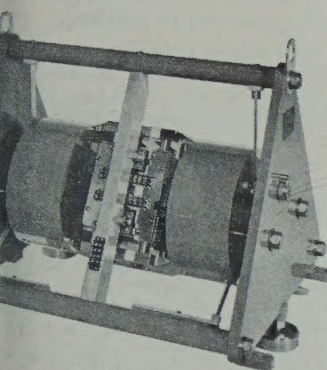
(An application form is also inserted in the closing pages of this issue.)

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers

SHORT-PERIOD SEISMOGRAPH INSTRUMENTS

BENIOFF SEISMOMETERS

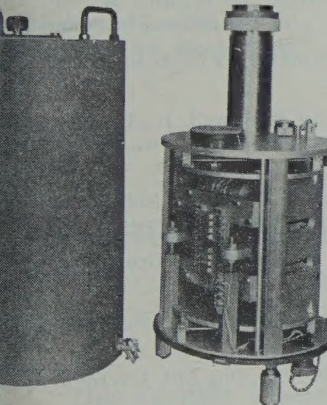
Vertical Model 1051 — Horizontal Model 1101



The exceptional magnification and yeoman service provided by these Benioff seismometers have earned them acclaim in seismological observatories throughout the world. Their natural period is 1 cps. The variable-reluctance, push-pull transducer has eight 125-ohm coils which can be connected in any series-parallel combination, providing flexibility of operation. A 237-pound steel inertial mass enables each coil to generate 1 mv output per micron of earth motion at 1 cps. Available magnification is over 1,000,000 when driving a Geotech short-period recording galvanometer without an amplifier, and is sufficient to meet any requirement in the quietest location in the world. Output power is also sufficient to drive more than one galvanometer, or to enable remote operation more than a mile from the recorders without an amplifier. The variable-reluctance design is inherently trouble-free—damping is electro-magnetic, and there are no close adjustment tolerances. Accurate calibration can be performed locally or remotely at operating gain.

BENIOFF PORTABLE SEISMOMETERS

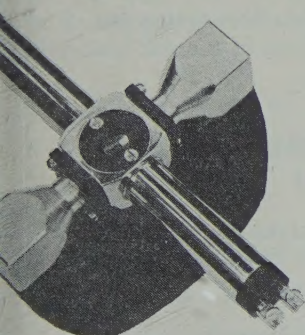
Vertical Model 4681 — Horizontal Model 6102



These are small, portable seismometers whose transducers are similar to the above. The rugged, 32-pound transducer also serves as the inertial mass, in place of the large mass of the standard Benioff seismometer. Each transducer coil generates 0.2 mv output per micron of earth motion at 1 cps. Available magnification is over 300,000 when driving a Geotech short-period galvanometer without an amplifier. This is sufficient for all but a very few locations in the world having the lowest of seismic background noise levels. Used with the Geotech phototube amplifier, useful gain is limited only by local seismic background noise. The seismometer can be furnished with four 20K ohm transducer coils if it is desired to use an electronic amplifier having a high input impedance. Useful frequency range is up to 100 cps. The seismometer is ideal for portable field use—a delta rod suspension and cylindrical frame design add to its ruggedness. Calibration can be performed locally or remotely. Dimensions are 22.5"H x 10.75" diameter. Weight is 63 pounds.

SHORT-PERIOD RECORDING GALVANOMETERS

Series 4100 — ()



These are rugged galvanometers designed to provide high sensitivity, stability, and adjustability. Sensitivity is up to 3×10^{-9} amps/mm at an optical lever of 1 meter. Natural frequencies are available from 1 to 50 cps. Horizontal and vertical spot position can be adjusted manually. The suspension frame is a separate, insulated, removable assembly. Rotation of either of the insulating end caps rotates the entire frame and suspension, and does not twist the suspension. Natural frequency can be set precisely, and can be varied in the field by sliding the nylon bridges which contact the gold suspension ribbon. A magnet is furnished with each galvanometer. A pedestal and galvanometer control (attenuator) are available on separate order.

For detailed specifications, price, or delivery, write to:

THE GEOTECHNICAL CORP.

SHILOH ROAD

GARLAND, TEXAS

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers

43RD ANNUAL MEETING
 AMERICAN GEOPHYSICAL UNION
 APRIL 25 through 28, 1962
 WASHINGTON, D. C.

*your participation
 is welcomed*

If you want to present a paper
 send the abstract BEFORE DEADLINE DATE to

<i>Hydrology</i>	<i>before January 2, 1962, to Harry E. Schwarz, 315 Ladson Road, Silver Spring, Md.</i>
<i>Meteorology</i>	<i>before January 9, 1962, to William Hiatt, US Weather Bureau, Washington 25, D. C. (abstract plus comprehensive summary)</i>
<i>Geodesy</i>	<i>before February 1, 1962, to Erwin Schmid, US Coast & Geodetic Survey, Washington 25, D. C.</i>
<i>Seismology</i>	<i>before February 1, 1962, to Benjamin F. Howell, Jr., Geophysical Laboratory, Mineral Sciences Bldg., Pennsylvania State Univ., University Park, Penna.</i>
<i>Oceanography</i>	<i>before February 1, 1962, to Arthur E. Maxwell, 5627 Potomac Avenue, N.W., Washington 16, D. C.</i>
<i>Volcanology, Geochemistry, and Petrology</i>	<i>before February 1, 1962, to David R. Wones, US Geological Survey, Washington 25, D. C.</i>
<i>Geomagnetism and Aeronomy</i>	<i>before February 1, 1962, to J. Hugh Nelson, US Coast and Geodetic Survey, Washington 25, D. C.</i>
<i>Tectonophysics</i>	<i>before February 1, 1962, to Mrs. Iris Borg, Box 579, Livermore, California.</i>
<i>Planetary Sciences</i>	<i>before February 1, 1962, to Robert Jastrow, Goddard Institute for Space Studies, 475 Riverside Dr., New York 28, N. Y.</i>
<i>General Program</i>	<i>to Leroy R. Alldredge, 10500 Royal Rd., Silver Spring, Md.</i>

Be sure to

send 3 copies, including ribbon copy, of the abstract
 double or triple space

make it concise—about 200 words

use correct format: John Doe (Physics Dept., Iowa State University, Ames, Iowa) Short-Period Micropulsations—The geographical extent,

see the December *Transactions* for complete details on the meeting

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers



McGRAW-HILL
BOOK
COMPANY

APPLIED CLAY MINERALOGY

By RALPH E. GRIM, University of Illinois. *McGraw-Hill Series in Earth Sciences.*
Available in January, 1962.

The primary object of this book is to analyze the fundamental structure and composition of clay materials in relation to the specific uses for them. As a companion volume to the author's *CLAY MINERALOGY*, this book will be a valuable reference work in geology and mineralogy, soil mechanics, ceramics, metallurgy, and petroleum engineering.

PHOTO GEOLOGY

By VICTOR C. MILLER, *Miller & Associates, Inc., Denver Colorado*, 249 pages,
\$13.50.

A text for graduate and advanced undergraduate courses in photogeology or photogeologic interpretation covering briefly the concepts of aerial photography and photogrammetric instrumentation, and then, in greater detail, stereoscopy and procedures. Of special interest is the final third of the text which contains stereo-pair aerial photographic illustrations accompanied by individual texts, exercises, and pertinent topographic or geologic maps, sketches and cross-sections.

330 West 42nd Street
New York 36, New York



SEND FOR
COPIES ON
APPROVAL

BULLETIN (IZVESTIYA), ACADEMY OF SCIENCES, U.S.S.R. GEOPHYSICS SERIES

Subscriptions for 1961 series now available

This monthly Russian publication, perhaps the leading journal of Geophysics of the U.S.S.R., is being translated and published in an English edition for the year 1961 by the American Geophysical Union. The twelve numbers in Russian cover about 2000 pages. Published with the aid of a grant from the National Science Foundation.

Send subscriptions now to

AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W.

Washington 5, D. C., U.S.A.

Subscription rates: \$25.00 for the volume of 12 numbers (\$20.00 to individual members of AGU subscribing for personal use)

Numbers will be mailed as issued.

The English edition of this publication for 1957 has been translated and published for the American Geophysical Union by Pergamon Press. This volume may be ordered through the American Geophysical Union at a price of \$25.00. The 1958, 1959, and 1960 series are available at a price of \$25.00 for each volume of 12 numbers. Titles and authors of the papers contained in the series have been published in recent issues of the *Transactions, AGU*.

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers

BREADTH

A new facility, a new scientific community, and new and ambitious scientific goals—this is the Sperry Rand Research Center. Here academic curiosity, scientific objectivity and true interdisciplinary exchange will combine to fill broad and varied basic research needs of the many divisions of Sperry Rand Corporation in the following areas: Materials Sciences including solids, mechanics, magnetics, and physical chemistry; Radiation Sciences including microwave physics, plasmas, infrared, optics, energy conversion, and communications; Applied Mathematics and Theoretical Physics; Earth and Life Sciences.

Our immediate needs for top scientific talent are in the areas of:

- Seismology
- Oceanography
- Geomagnetism
- Atmospheric Physics

If you are interested in fundamental research in any of the above fields, direct your communication in confidence to:

Frederick M. Swope, Jr.,
Sperry Rand
Research Center,
North Road,
Sudbury, Massachusetts

an
important
word
at the
new

SPERRY RAND RESEARCH CENTER
SUDBURY, MASSACHUSETTS

Journal of GEOPHYSICAL RESEARCH

ME 66

OCTOBER 1961

No. 10

Variations of the Cosmic Radiation in November 1960

J. A. LOCKWOOD AND M. A. SHEA

*University of New Hampshire
Durham, New Hampshire*

Abstract. Two increases of about 100 per cent in the cosmic-ray nucleonic intensity were recorded at Mount Washington and Durham, New Hampshire, on November 12 and 15 following large solar flares. A third small increase was observed on November 20. The two large increases exhibited different features, and these differences are discussed in relation to solar phenomena. From a study of the world-wide nucleonic intensity data it is found that: (1) the rigidity spectrum can be expressed as KP^{-n} , with $n=6$ for both flares; (2) for the increase on November 12, there is particle dispersion, and an approximate evaluation of the diffusion has been made; (3) the classical impact-zone theory does not apply for these flares, evidence being presented for the existence of magnetic fields and some scattering effects between the earth and the sun.

INTRODUCTION

Two large increases and at least two rapid push decreases in the cosmic radiation at the earth were observed between November 12 and 1960. A third, smaller increase was observed November 20. The increases in the cosmic-ray intensity occurred shortly after class 3+ solar flares in region HH25 [*Preliminary Report on Solar Activity*, 1960]. The Forbush decreases showed large geomagnetic disturbances and indicate the immersion of the earth in solar gas clouds or beams.

This unusual sequence of events is important in the study of the electromagnetic state of the earth-sun region, because we have the injection of relativistic particles from the sun into a region where solar-controlled modulating mechanisms were operative to suppress the galactic cosmic radiation. The difference in the shapes and magnitudes of the two large increases, representing the injection of particles, suggests that the electromagnetic conditions in the earth-sun region were different.

We report here: (1) the intensities of the nucleonic component of cosmic radiation at Mount Washington and Durham, New Hamp-

shire; (2) the rigidity spectra of the relativistic solar-produced particles; (3) calculations and interpretation of the preferred directions for the solar particles using data from the world-wide network of cosmic-ray stations; (4) evidence for and an interpretation of the dispersion effects during the first increase.

COSMIC-RAY VARIATIONS AT MOUNT WASHINGTON AND DURHAM

The hourly average nucleonic intensities at Mount Washington (elevation, 1910 meters) and Durham (sea level) for November 11-16 are shown in Figure 1. The differences in shape and magnitude of the two events are evident and appear to be characteristic of all nucleonic detectors recording the increases.

A more detailed comparison of the two events can be made from Figures 2 and 3, where the 10-minute average counting rates at Mount Washington are plotted. In Figure 2 we see that the maximum at 1614-1624 was followed by a slight decrease, and then a sudden rise at 1900. The rapid dip in the intensity at 1920-1930 represented the effects of a Forbush-type decrease, recorded more strikingly by detectors

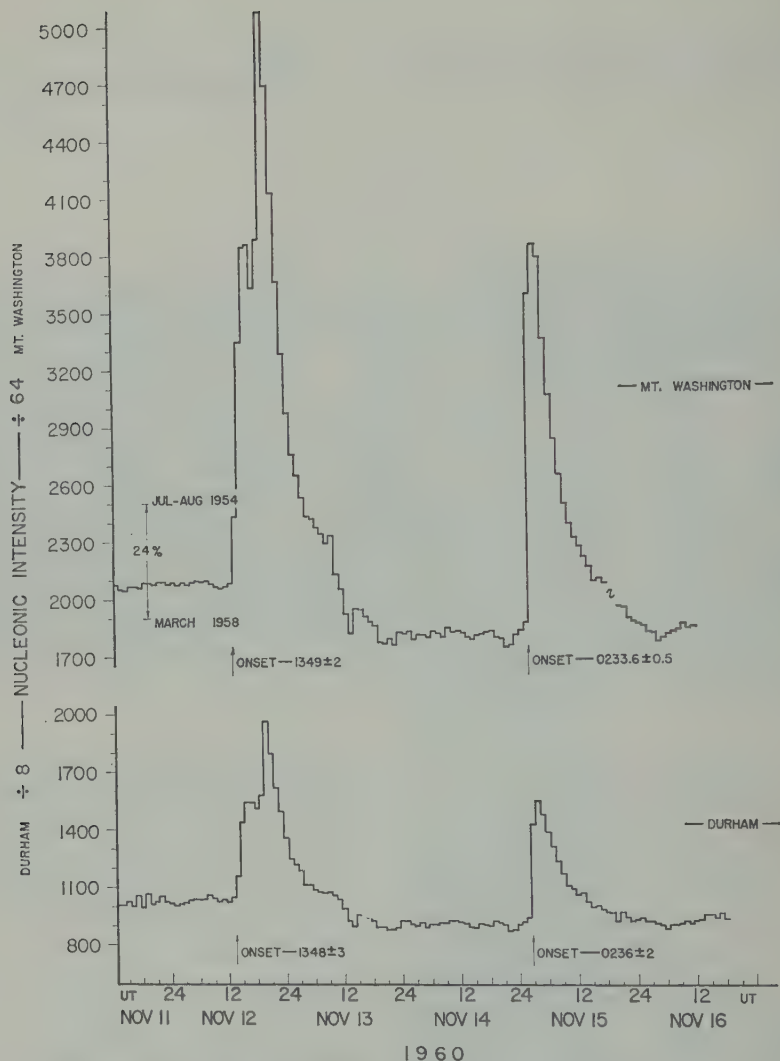


Fig. 1. Hourly average nucleonic intensity at Mount Washington and Durham. All readings have been corrected for barometric pressure changes. The onset times are indicated for reference. The nucleonic intensities at Mount Washington in 1954 and 1958 are shown.

with a mean response at a higher rigidity. The Forbush decrease between the two flares, commencing at 1030 on November 13, reduced the level at Mount Washington 14 per cent and at Durham 13 per cent below the preflare level. The percentage increase recorded at each station for the second flare was calculated from this depressed intensity.

To determine precisely the onset time for the second flare, we have used the 10-minute running averages, as shown in the insert of Figure 3.

Such a procedure was necessary because so many difficulties were encountered in the recording system at Mount Washington from 2100 on November 14 to 0232 on November 15. The zero line was taken as the average intensity 1200–2400 on November 14. The rapid oscillations between 0315 and 0500 were much greater than the statistical fluctuations, and therefore do not represent real variations in the cosmic radiation [Steljes, Carmichael, and McCracken, 1961].

A third increase in the cosmic radiation

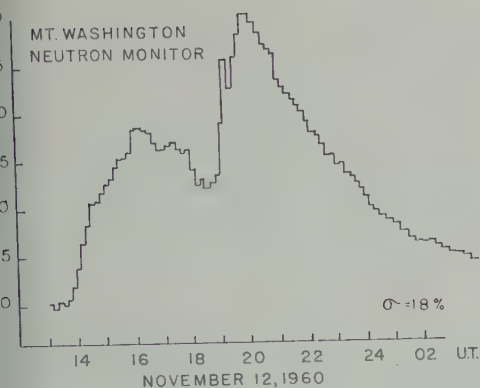


Fig. 2. The 10-minute average nucleonic intensity at Mount Washington during the first flare event, normalized to the average intensity from 12 November 11 to 12 on November 12.

recorded on November 20 following a west-limb flare, also believed to have occurred in the region HH25 [Carmichael, Steljes, Rose, and Wilson, 1961]. The increase observed at Mount Washington is shown in Figure 4. The data here are conclusive only when supplemented with observations from other stations.

During the decay phase of the first flare effect unusual and large variations in the nucleonic intensity occurred, shown in Figure 5. There was a large and sudden Forbush decrease commencing at 1030 on November 13, the minimum being reached about $3\frac{1}{2}$ hours later. Until the start of the Forbush decrease, the intensity had declined smoothly. The 5 per cent increase just before the Forbush event is suggestive of an albedo effect, produced by the outward-moving solar gas cloud which reduced the galactic cosmic-ray flux after 1030. Data from other stations do not make such an interpretation unique.

Data from these flare effects, as well as from events on February 23, 1956, and May 4, 1960, are summarized in Table 1. It appears that the time of the cosmic-ray effect is most rapid when the flare occurs on the west limb [McCracken and Palmeira, 1960].

RIGIDITY SPECTRUM OF FLARE-PRODUCED PARTICLES

We can calculate the rigidity spectrum of the solar-produced radiation by the following method. Let us assume the response of a nucleonic or

mesonic detector during the flare-associated increase to be given by [Webber and Quenby, 1959]:

$$N_v(P, x, t) = \int_{P_0}^{\infty} S(P, x) j(P, t) dP \quad (1)$$

where P_0 is the vertical cutoff rigidity for the detector, x is the depth of the atmosphere, $j(P, t)$ is the differential primary rigidity spectrum in the vertical direction at a particular time, and $S(P, x)$ is the response of the detector due to particles arriving vertically at the top of the atmosphere. In the actual calculation, the integral in equation 1 was replaced by a summation process. We will use the counting rates observed at Mount Washington, Durham, and the Massachusetts Institute of Technology [Steljes, Carmichael, and McCracken, 1961] to determine the differential primary rigidity spectrum. Assuming that the intensity of radiation in the asymptotic cones of acceptance [McCracken, 1961] is essentially the same, we can use their difference in response to determine this spectrum. A small correction was made for the difference in cutoff rigidity at Mount Washington and Durham. Since the meson detector does not respond to particles with $P \sim 1.5$ BV, no correction was necessary.

If we assume that the rigidity spectrum was of the form $KP^{-\alpha}$ and that only protons were present (as in equation 1), we can use the specific yield

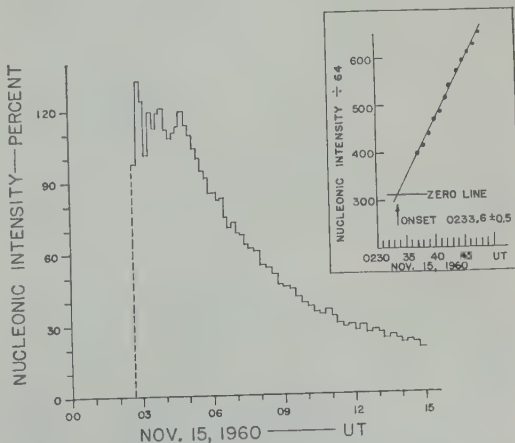


Fig. 3. The 10-minute average nucleonic intensity at Mount Washington for the second flare effect, normalized to the average intensity from 12 to 24 on November 14. The 10-minute running averages to determine the onset time are shown in the insert.

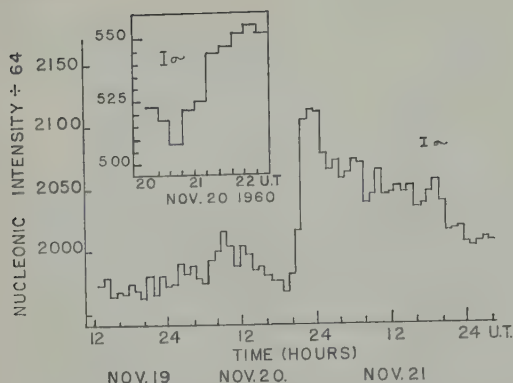


Fig. 4. Hourly average nucleonic intensity at Mount Washington for the increase on November 20. The 15-minute intensities near the onset time are shown in the insert.

functions of Webber and Quenby [1959] to find n and K . Since the specific yield function for particles with rigidity less than 2 BV is not known for nucleonic detectors, the curves given by Webber and Quenby had to be extrapolated to 1 BV. The large number of low-rigidity particles present during a flare effect make the calculation of n extremely dependent upon the specific yield function for primary particles of 1 to 2 BV. Since Mount Washington is at an altitude of 820 g cm⁻², the specific yield functions for this station were determined by interpolation of the yield curves for different altitudes. We took the quiescent primary rigidity spectrum just before the flare to be different from 1954, using a modulation factor of $1.5/P$ [Lockwood, 1960]. To increase the reliability of the estimates, the spectrum was divided into intervals of 0.5 BV, for the summation process. From the ratios of Mount Washington to Durham and Mount Washington to MIT the following results were obtained:

In the first flare on November 12, $n = 4$ for the differential rigidity spectrum from onset to 1500. From 1500 to 1830, $n = 6$. In both cases the spectrum was cut off for $P \geq 7$ BV, consistent with the small increase observed by the MIT detector and the absence of an increase for nucleonic detectors with $P_c \geq 9$ BV. From the magnitude of the increase in nucleonic intensity at Mount Washington, we can estimate the flux of protons. Hence, we may write for the differential spectrum of protons between 1600 and 1630 on November 12

$$j(P) = 5 \times 10^5 P^{-6} \quad 1 \leq P \leq 7 \text{ BV}$$

where $j(P)$ is the differential flux arriving vertically in particles m⁻² sec⁻¹ ster⁻¹ (BV⁻⁶). For the second flare effect on November 20, the same method gives $n = 6$, with a maximum rigidity of $P = 6$ BV at 0400 UT. No increase was recorded at MIT for this event. From the maximum intensity recorded at Mount Washington, the differential flux was

$$j(P) = 10^6 P^{-6} \quad 1 \leq P \leq 6 \text{ BV}$$

The error in the constant factor is about 10 per cent.

From an inspection of the increases observed at other nucleonic detectors, we find confirmation of these results for the rigidity spectrum. First, the nucleonic detectors at Chicago (Simson, private communication) and at Deep River [Steljes, Carmichael, and McCracken, 1960] recorded the same magnitude of increase between 1330 and 1415 UT on November 12. This might have been due to absence of particles with rigidities ≤ 1.5 BV, the vertical cutoff rigidity for Chicago, provided that the asymptotic cone of acceptance are approximately the same at the two stations. This appears to be a reasonable assumption. If particles with rigidities less than 1.5 BV were absent, the best value of the exponent of the rigidity spectrum would still be $n = 4$, based upon the calculated ratio of Mount Washington to Durham.

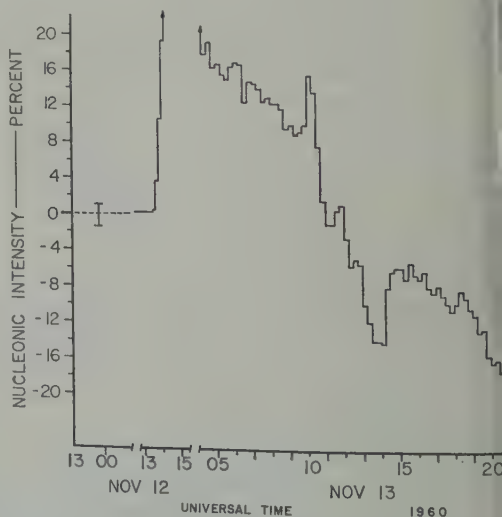


Fig. 5. The 15-minute average nucleonic intensity at Mount Washington during the decay phase of the first flare effect, normalized as in Figure 2.

TABLE 1. Summary of Cosmic-Ray Variations at Mount Washington and Durham Including Data from February 23, 1956, and May 4, 1960, Events

Date	Solar Flare*				Cosmic-Ray Increases			Per Cent Increase		$\Delta t \dagger$
	Location	Onset	Max.	End, UT	Onset	Max.	End, UT	Mt. Wash.	Durham	
2/60	1°W, 27°N	<1323	1426	1922	1349 \pm 2	1618 (2002)§	0940	85 (152)	50† (91)†	6 h
5/60	33°W, 26°N	0207	0221	0427	0233.6 \pm 00.5	0256	~2330	131	70†	22 m
4/60	90°W, 10°N	~1015	...	1025	1031	1041	1200	214	140	10 m
23/56	85°W, 25°N	0331	0342	1415	0350	0415	~2000	...	3000†	25 m

Solar data from CRPL Reports, National Bureau of Standards, Boulder, Colorado.

Time for cosmic-ray intensity to reach maximum.

Based upon hourly average counting rates, other percentages upon 5-minute averages.

Second and largest peak intensity.

From 1415 to 1730 it appears that the flux of particles with $P \leq 1.5$ BV increased. After 1720 was assumed that the cutoff rigidity at Durham increased to about 1 BV. Magnetic data from Fredericksburg showed a s.c. magnetic storm 1350 UT, November 12, but the geomagnetic field remained constant to $\sim 20 \gamma$ from 1445 to 10, after which large variations occurred. Changes in cutoff rigidities after 1730 must be considered in evaluating the solar particle spectrum.

Second, the increases in nucleonic intensity recorded at Leeds (Marsden, private communication) and Uppsala (Sandstrom, private communication) support these rigidity spectra. Since these stations are at sea level and the vertical cutoff rigidities are known, the relative intensities expected at the two stations can be calculated. For the first maximum at 1600-1630 on November 12, the observed ratio Uppsala/Leeds was 2.0. If $n = 6$, the calculated ratio for these stations is 1.9, in agreement with the observed value. At 1810 the ratio of these stations was about unity, consistent with the same magnetic cutoff rigidity at both stations. For the second flare event, the ratio of Uppsala to Leeds was 2.0, consistent with $n = 6$.

Third, the data from nucleonic detectors at Zugspitze (Meyer, private communication), Pic du Midi (Freon, private communication), and London (Thambyahpillai, private communication), if combined with the Leeds and Uppsala data, support $n = 6$ for both flares at the times indicated previously.

IMPACT-ZONE PHENOMENA

Firor [1954] has calculated the impact zones on the earth for relativistic particles from the sun and applied these to flare events prior to 1956. Lüst and Simpson [1957] calculated additional orbits of charged particles in the geomagnetic field and used them to discuss the impact zones for the February 23, 1956, flare effect. Some discrepancies between the observed and calculated increases from impact-zone theory are observed, particularly for polar stations such as Thule and Resolute. In this section we will consider the results of the November and May 1960 cosmic-ray increases using the conventional impact-zone theory, i.e., Firor's method, and the method of asymptotic cones of acceptance [McCracken, 1961].

Conventional impact-zone theory. Essentially Firor's [1954] method consists of determining the impact points at the earth of orbits of protons of different rigidities which have their origin at the sun.

With a range of rigidities and a finite source size, these discrete points merge into zones, occurring at roughly 0900, 0400, and 2000 local time when the source is at local noon. For the flare of November 12, we have determined the impact zones at the earth as shown in Figure 6. The following assumptions were made: a flat spectrum of protons from 1 to 6 BV; a source diameter of 40° in the direction of the sun. The onset times are not sharply defined because the rate of rise for the effect was slow. The majority

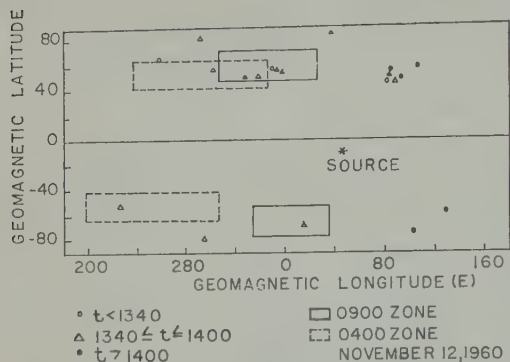


Fig. 6. The impact zones for the cosmic-ray increase on November 12 from conventional impact-zone theory. The source is taken in the direction of the sun.

of stations with early onset times, except for the polar stations, lie in the impact zones. At later times all stations with $P_c \leq 5$ BV recorded increases, indicative of some isotropy. The particle dispersion present, which will be discussed, introduces a further complication. Stations such as Jungfraujoch and Pic du Midi may have an early onset because higher-rigidity particles arrived first. A detailed examination shows that the source size can be reduced to 20° in diameter with no essential change in the results.

The same method was applied to the increase on November 15. The results are presented in Figure 7. The stations observing an early onset, especially in the southern hemisphere, clearly do not lie in the impact zones. The early onset times for polar stations cannot be explained on the basis of such calculations. Better agreement can be obtained by shifting a source of the same size approximately 70° to the west, but still the effects at the polar stations are unexplained.

Preferred directions. In this method, proposed by McCracken [1961] and based upon the work of Brunberg and Dattner [1954], the direction of the sky which is scanned by a nucleonic detector at a given place on the earth is determined. This direction, found by evaluating the deflection of the particles in the geomagnetic field, is expressed by the direction of an orbit of a particle far away from the earth. This is called the asymptotic direction. The direction of the velocity vector may be defined by latitude and longitude angles, measured with respect to the

detector location, and hence provides information on the regions of space from which particles of different rigidities, azimuth, and zenith angle come. Detectors at different locations on earth scan different regions of space, although with nucleonic detectors these regions may be extensive because of the large longitude spread of particles arriving near cutoff rigidities. The expected increases at the earth are then calculated when a source of particles has a specific location in space, as at the sun.

The details of these calculations, which are in the process of publication by K. G. McCracken, indicate that the source of the extra radiation observed on November 15 should be shifted approximately 70° W to produce the observed increases. From the increases observed at Mawson, McMurdo, and Thule at 0300 and 0310, it appears that the radiation was becoming isotropic, and by 0345 all stations with the same cutoff rigidities recorded the same magnitude of increase. It is important to note that McMurdo and Thule recorded early onsets of a small increase even though the asymptotic direction of these stations are $>90^\circ$ with respect to the shifted source. This indicates that some scattering of the flare particles occurred in the earth-sun region, especially in the immediate vicinity of the earth.

The increase on November 12 was not analyzed by this method because the radiation was not anisotropic, and consequently the increase did not lend itself to analysis by this method. A similar analysis, for example, of the May

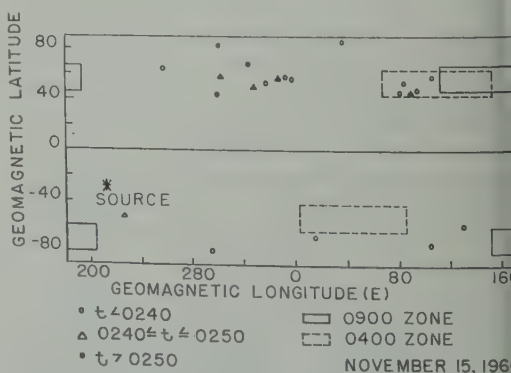


Fig. 7. The impact zones for the cosmic-ray increase on November 15 from conventional impact-zone theory. The source is taken in the direction of the sun.

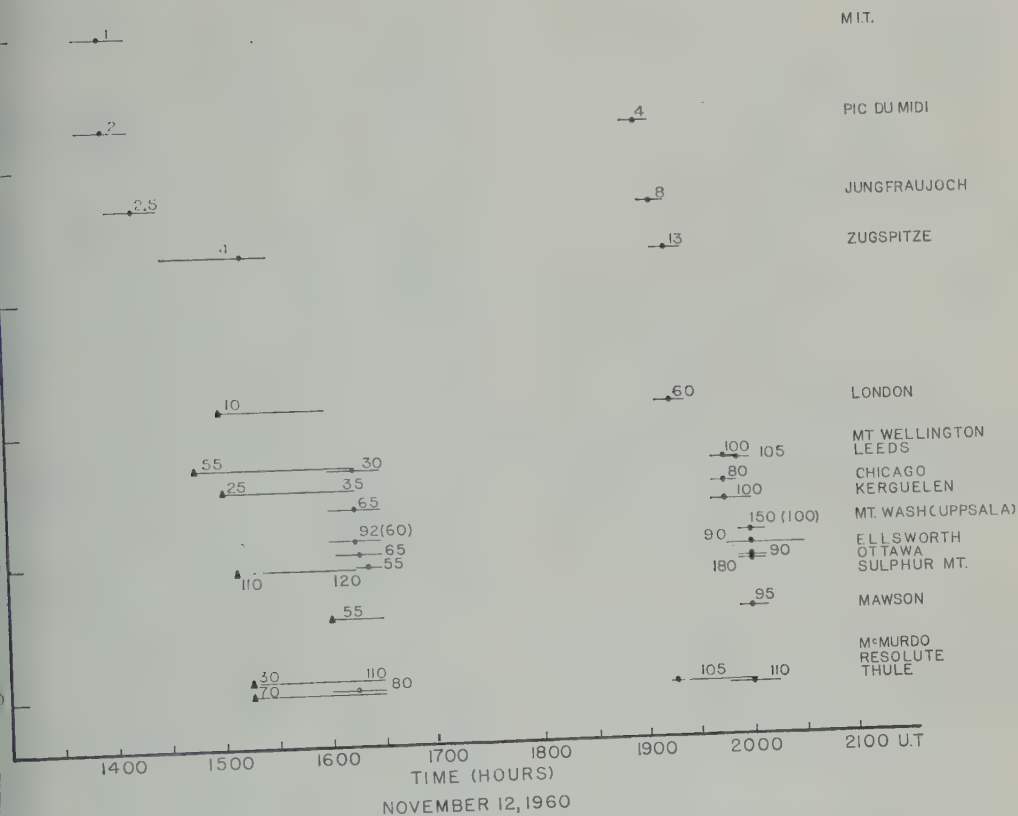


Fig. 8. The time of maximum intensity for nucleonic detectors with different cutoff rigidities for the first flare event. A circle indicates two well-defined maxima at ~ 1600 and ~ 2000 UT. A triangle is used for stations with a broad, variable maximum between 1445 and 1800 UT. The magnitudes are shown for each station.

30, event [Trainor, Shea, and Lockwood, 1960] shows that the source must be shifted about 50° east to account for the observed increases, in agreement with the results of McCracken [1961]. However, using conventional impact zones, a westward shift of the source of 90° is necessary. No significance can be attached to the different magnitudes of the shift by the two methods because different approximations to the geomagnetic field and cutoff rigidity at each station were used.

ENERGY DISPERSION EFFECTS

The difference in shape of the cosmic-ray decreases on November 12 and 15 is evidence for changes in the electromagnetic conditions between the earth and the sun. The slow rate of rise on November 12 suggests that the cosmic

rays diffused slowly to the earth. Such diffusion processes should be energy-dependent. The earlier maximum at the MIT meson detector indicates energy dispersion effects. No significant differences in onset times were observed, because the initial rate of rise was too slow. The impact zones for this event were not clearly defined, indicative of considerable scattering or a large source. In contrast, the event on November 15 exhibits a rapid rise to full intensity with little evidence for energy dispersion and definite preferred regions on earth for greatest intensity. Such differences have been discussed for earlier flare events by *McCracken and Palmeira* [1960], and a general model has been proposed to account for them.

Figure 8, which shows the time of maximum for nucleonic detectors with different cutoff

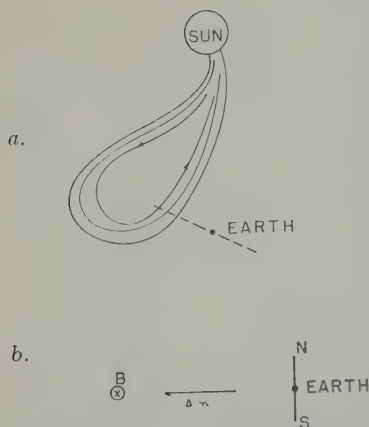


Fig. 9. (a) A representation of the relationship of sun, earth, and solar plasma at about 1430 UT on November 12. (b) Schematic drawing of the earth, component of the particle density gradient, and the magnetic field in the solar plasma at 1630 UT, viewed perpendicular to Figure 9a.

rigidities, illustrates this energy dispersion. The differences in times and magnitudes of the maxima are attributed to the change in the rigidity spectrum about 1500 and the dependence of the specific yield upon rigidity.

INTERPRETATION OF DIFFERENCES IN NOVEMBER 12 AND 15 INCREASES

Steljes, Carmichael, and McCracken [1961] have proposed a model to account for the differences in the two solar-flare events. Such a model is an extension of the ideas of *Cocconi, Greisen, Morrison, Gold, and Hayakawa* [1958], *Piddington* [1958], and *McCracken and Palmeira* [1960]. The interpretation presented here is based on essentially the same model. We can picture the earth, sun, and the plasma cloud, which arrived at the earth about 1830 UT to produce the s.c. geomagnetic storm, as shown in Figure 9a. This gas cloud or beam probably originated in a solar flare at 0304 November 11 [*Steljes, Carmichael, and McCracken*, 1961]. At 1430 UT the earth was located outside this cloud, in which the solar cosmic rays were magnetically trapped. The sudden increase of intensity at 1900 must then represent the earth moving into this region sufficiently far to sample the large intensity of low-energy particles. Before this time the effects at the earth had been from particles leaking out of the magnetic bottle. The dip in cosmic-ray intensity at 1920–1930 represents the decrease

in the galactic radiation, the earth now being inside more than about 2 gyroradii for galactic cosmic rays of rigidities $P > 2$ BV [*Steljes, Carmichael, and McCracken*, 1961].

The slow rate of rise and the energy dispersion for this event were then due to the slow, energy-dependent rate of diffusion of the trapped solar particles out of this region. This is in marked contrast to the event on November 15, when the earth was presumably inside the magnetic trapping region. Now, for the first event, the diffusion of these particles must effectively occur across the magnetic field lines of the confinement region. Let us attempt to estimate the diffusion rates involved in such a process.

First, we can calculate the magnitude of the increase that would have been recorded at Mount Washington had the earth been inside the trapping region on November 12. The onset of the flare from visual observations was 1322 UT. Hence, inside the magnetic bottle the onset of the cosmic-ray effects would have been at $t_0 \sim 1340$ UT. This 18-minute delay is comparable to that observed for the events of February 23, 1956 [*Meyer, Parker, and Simpson*, 1956; *Lüst and Simpson*, 1957], November 13, 1960, and May 4, 1960. With similar rates of rise, the time of maximum intensity t_m would have been at ~ 1400 UT. Since Mount Washington observed a 1.5x increase at 2000 UT, when the earth was inside the magnetic trapping region, the intensity inside at t_m would have been 124x. We have assumed a $t^{-1/2}$ dependence of intensity upon time [*Meyer, Parker, and Simpson*, 1956], the same as observed for the decay of the November 15, 1960, event. For the Durham detector the extrapolated intensity would have been 69x, about twice the actual increase observed in February 23, 1956 [*Lockwood, Yings, Calawa, and Sarmaniote*, 1956]. In the above cases, x is defined as the pre-flare level of intensity.

At 1430 UT, the Mount Washington intensity was 50 per cent above normal. The ratio of intensity, R , outside to that inside, can be used to estimate the diffusion coefficient, D , as follows. Let I be the intensity outside at time t and distance r from the source (or inside region); then

$$I = I_m \left(\frac{t}{t_m} \right)^{-3/2} \exp \left(-\frac{r^2}{\pi D t} \right) \quad (1)$$

[*Meyer, Parker, and Simpson*, 1956]. Assuming

0.2 AU, based upon the time delay of the magnetic effects, with $t = 50$ minutes, after 20 minutes, we find $D \sim 3 \times 10^{20}$ cm²/sec. Assuming that D_{\perp} represents diffusion across magnetic lines, it can be related to the diffusion parallel to the field by

$$D_{\perp} = D_{\parallel} / (1 + \omega^2 \tau^2) \quad (3)$$

(Schiff, 1956], where ω is the gyrofrequency in a field of strength B , τ is the collision time, and $\omega = \frac{1}{2} l_e v$. (Relativistic expressions for ω and τ are used.) Let us assume the collision length for all-angle scattering $l_e \sim 0.1$ AU, which is smaller than the gyroradius for these particles. D_{\parallel} is 1.5×10^{22} cm²/sec, a value comparable to that used by Meyer, Parker, and Simpson [1956]. If we consider 2-BV protons with the average magnetic field $\sim 10^{-4}$ gauss, then D_{\perp} is $\sim 4 \times 10^{19}$ cm²/sec. Therefore, for field strengths of the order of the average value measured recently [Coleman, Bennett, Judge, and Smith, 1960], the diffusion across field lines is a reasonable interpretation.

For the data recorded simultaneously at MIT, we assume the response to be from ~ 5 -BV particles and that the rigidity spectrum is given by KP^{-n} with $n \sim 4$. MIT recorded a maximum increase of ≤ 1 per cent. Thus, D from equation (3) is $\sim 2.5 \times 10^{20}$. This estimate may be incorrect, because the threshold response for the MIT scintillation detector is not precisely known. The corresponding value of D_{\perp} for $B = 10^{-4}$ gauss is 1.5×10^{20} cm²/sec, with D_{\parallel} independent of rigidity.

We can check these estimates, since the time delay between the maximum at Mount Washington and the maximum inside the beam was ~ 2 hours. Assuming that this maximum represents a large flux of 1 BV protons, we would expect

$$r^2 \sim \pi D t \quad \text{where } t \sim 2 \text{ hr}$$

Thus, $D \sim 1 \times 10^{20}$ cm²/sec. For a 1-BV particle in a field $B = 10^{-4}$ gauss, with $l_e = 0.1$ AU, $D_{\perp} \sim 3 \times 10^{19}$ cm²/sec.

Two other features of the increase on November 12 support this interpretation. First, the decay phase at Mount Washington from 1630 to 1830 UT was not as $t^{-3/2}$ but as an exponential time. If we assume the earth about 0.1 AU from the edge of magnetic trapping region, the radiation approximately isotropic, and $t_0 = 1400$ T, we can compare the observed decay with

that calculated from equation 3 considering only the exponential term. If $D \sim 3 \times 10^{20}$ cm²/sec for 2-BV particles, the calculated counting rate at Mount Washington at 1820 UT is $0.70 I_{\text{max}}$. The observed intensity was $0.66 I_{\text{max}}$.

Second, there is an observed anisotropy at 1630 UT as measured by the stations at McMurdo and Thule [Pomerantz, Duggal, and Nagashima, 1961]. The ratio of McMurdo to Thule is ~ 1.20 at this time. This may be attributed to the effect of a gradient in particle density within a magnetic field. Referring to Figure 9b, we see that a component of the gradient in density of cosmic-ray particles with the magnetic field oriented as shown produces a greater influx of particles at the south than at the north pole of the earth. Dattner and Venkatesan [1959] and Elliot [1960] have accounted for the daily variation on the basis of such a mechanism. Let us assume that the fractional difference of intensity is

$$\Delta I/I \cong K \rho$$

where $K = \Delta n/n/\Delta r$, n is related to particle density, and ρ is the gyroradius. K also equals $-1/L$, where L is the diffusion length [Dattner and Venkatesan, 1959]. For 2-BV particles in an average field of 5×10^{-5} gauss, $\rho = 1 \times 10^{11}$ cm. Hence, $K = 1.5 \times 10^{-12}$ cm⁻¹, if $\Delta I/I = 20$ per cent. Now

$$K = 1/L = 1/\sqrt{D\tau}$$

where D is the diffusion coefficient and l_e the time between collisions. Using $D = 3 \times 10^{20}$ cm²/sec and $l_e \sim 0.1$ AU, we find $K \sim 8 \times 10^{-12}$ cm⁻¹. Although this agrees within an order of magnitude, it should be noted, first, that there are gradients in the magnetic field, which are important but are difficult to evaluate; and second, that some of the difference between McMurdo and Thule might have been produced by differences in the region of space scanned by each detector [Brunberg and Dattner, 1954]. However, this agreement suggests that some of the anisotropy was produced by the large gradient of cosmic-ray density in the presence of the magnetic field of the trapping region.

The much greater increase recorded at College, Alaska (Korff, private communication), indicates that there was a longitude dependence at ~ 1600 UT. At this time College was looking in the direction of the sun. Part of the difference between McMurdo and Thule may also be

explained on this basis, since the effective direction for McMurdo was toward the sun.

CONCLUSIONS

From this investigation of the solar-flare-associated cosmic-ray increases on November 12 and 15, we conclude the following:

1. The rigidity spectra for the two large cosmic-ray increases were:

Nov. 12		
1430-1500	$j(P) = KP^{-4}$	$1 \leq P \leq 7 \text{ BV}$
1600-1630	$j(P) = 5 \times 10^6 P^{-6}$	$1 \leq P \leq 7 \text{ BV}$
Nov. 15		
0400	$j(P) = 10^6 P^{-6}$	$1 \leq P \leq 6 \text{ BV}$

The change in rigidity spectra for the first event is the result of the energy dependence of the diffusion process.

2. From the preferred directions or conventional impact zones, it is concluded that the source of the particles on November 15 must be shifted 70°W of the sun. The magnitude can be evaluated better from the first method. Such a shift could be produced if the particles were to travel along field lines from the sun to the earth, as shown in Figure 9a. A similar shift of about 50°W is necessary to account for the observed increases on May 4, 1960 [McCracken, 1961].

3. Since McMurdo and Thule recorded early onsets of a small increase on November 15 when the asymptotic directions of these stations were ~90° with respect to the shifted source, scattering of the solar particles must have occurred. Since the magnetic field decreases as we go out from the sun toward the earth in the magnetic cone, any uniform distribution of pitch angles $\leq \pi/2$ for the particles at the sun become collimated with pitch angles $\leq 1^\circ$ at the earth unless scattering occurs.

4. For the first event there is definite evidence of injection of solar particles within a magnetic cone, with slow leakage of particles out to the earth in the early phases of the event. Inside the cone the galactic cosmic radiation is suppressed [Cocconi, Greisen, Morrison, Gold and Hayakawa, 1958; Piddington, 1958; and McCracken and Palmeira, 1960].

The intensity within the magnetic cloud about 20 minutes after the onset of the flare was greater than on February 23, 1956, as extrapolated from the Mount Washington and Durham nucleonic intensities. The rapid Forbush

decrease at 1930 UT (see Fig. 2) only temporarily reduced the intensity at stations such as Mount Washington, which were moving into the high flux densities of solar protons, 1 to 3 BV rigidities.

5. The energy dispersion evident on November 12 has been interpreted as the result of diffusion across magnetic lines of force. This interpretation accounts for the slow rate of rise, differences in the time of maximum, the decay phase from 1630 to 1830 as observed at Mount Washington and the anisotropy at the two polar stations, McMurdo and Thule. The calculations are consistent with an average B field in the boundary of the gas cloud of $\sim 5 \times 10^{-5} - 10^{-4}$ gauss. The collision length of ~ 0.1 AU provides the field with an over-all regularity, the irregularities being less than the gyroradius of particles of 2 to 10 BV rigidity. The magnitude of the event on November 12 was greater than that of the event on February 23, 1956, but less rich in particles with $P > 5$ BV, and the earth was not in the proper position to receive the full impact. The storage times for the particles within the gas cloud are reasonable [Piddington, 1958; Meyer, Parker, and Simpson, 1956] for the diffusion coefficient estimated. The recent results of Winckler, Bhavsar, Masley, and May [1961] for the September 3, 1960, solar-flare event indicate a delay time of $2\frac{1}{2}$ hours between flare onset by optical observations and onset of cosmic-ray effects at Deep River or Mount Washington. Since this was an east-limb flare and the earth was probably situated far from the edge of the magnetic bottle, the solar particles must have diffused a considerable distance. Such a delay time is consistent with a distance, $r \sim 0.6$ AU, or a magnetic cone with a half-angle of about 30°.

Acknowledgments. The computational assistance of Mrs. Maurine Spiller is appreciated. The Mount Washington Observatory staff has conscientiously maintained the nucleonic detector there. Discussions with Mr. James Trainor and Dr. Kenneth McCracken were very helpful. Cosmic-ray data provided by the following are most gratefully acknowledged: R. R. Brown (Berkeley); H. Carmichael (Deep River); R. L. Chasson (Lincoln); I. Escobar (Chacaltaya); A. Freon (Pic du Midi, Port aux Francais); H. G. Houtermans (Jungfraujoch); S. A. Korff (College); P. L. Marsden (Leeds); B. Meyer (Zugspitze); Y. Miyazaki (Mt. Norikura); N. R. Parsons (Lae, Hobart, Mt. Wellington, Mawson); D. C. Rose (Ottawa, Churchill, Resolute, Sulphur Mountain); A. E. Sandström (Uppsala); O. Santochi (Ellsworth);

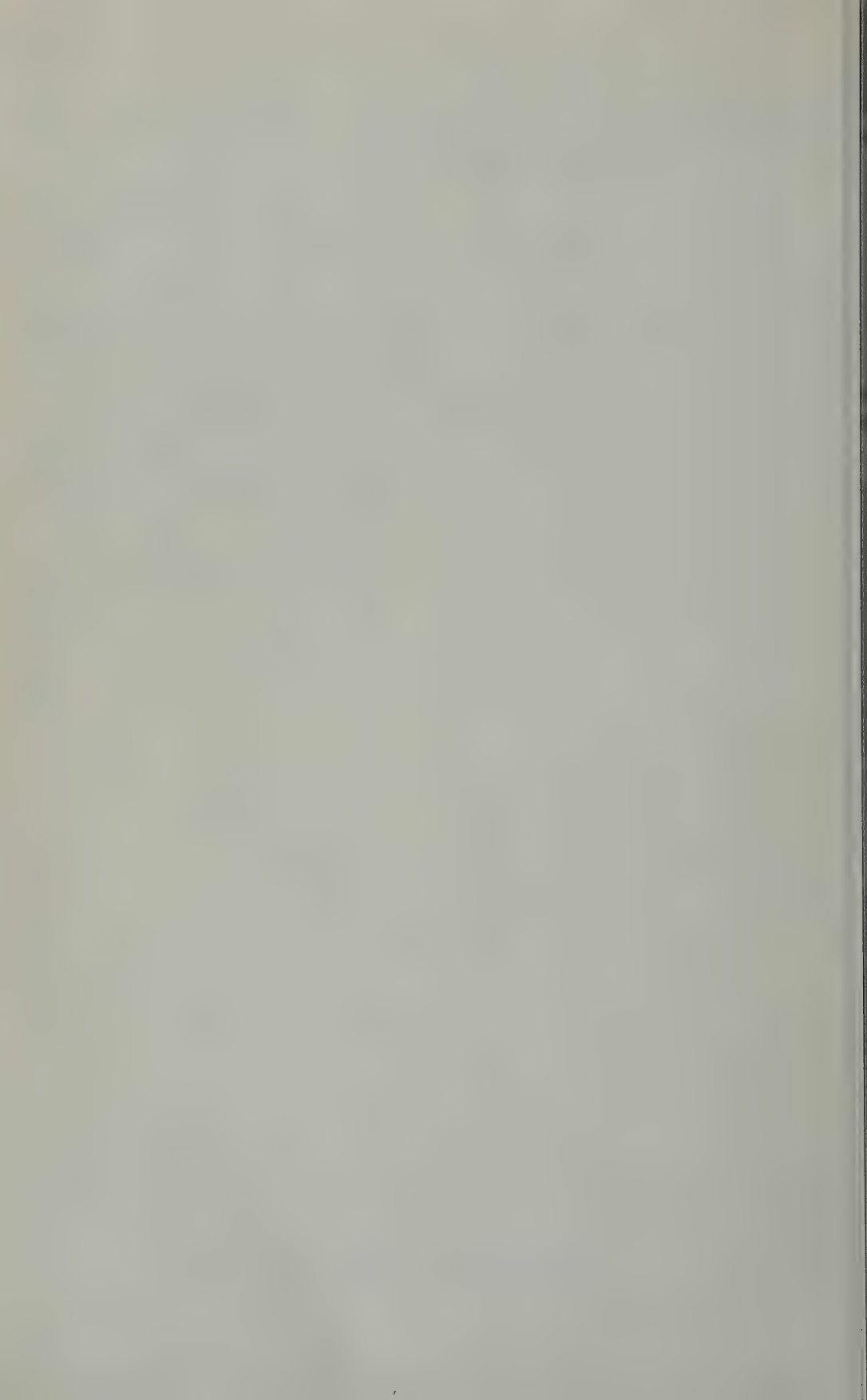
Schwachheim (Rio de Janeiro); J. A. Simpson (Chicago).

This research was supported by the National Science Foundation and the Geophysical Research Directorate, Air Force Cambridge Research Center.

REFERENCES

- Ellis, W. P., Motion of ions and electrons, *Handbuch der Physik*, 21, 383, Springer-Verlag, Berlin, 1956.
- Funkenberg, E. A., and A. Dattner, *Tellus*, 6, 254, 1954.
- Garman, H., J. F. Steljes, D. C. Rose, and B. G. Wilson, *Phys. Rev. Letters*, 6, 49, 1961.
- Goehner, G., K. Greisen, P. Morrison, T. Gold, and S. Hayakawa, *Nuovo cimento, Suppl.*, 8, 161, 1958.
- Holmes, P. J., Jr., C. P. Sonett, D. L. Judge, and E. J. Smith, *J. Geophys. Research*, 65, 1856, 1960.
- Dattner, A., and D. Venkatesan, *Tellus*, 11, 239, 1959.
- Elliot, H., *Phil. Mag.*, 5, 601, 1960.
- Forster, J., *Phys. Rev.*, 94, 1017, 1954.
- Lockwood, J. A., *J. Geophys. Research*, 65, 3859, 1960.
- Lockwood, J. A., H. E. Yingst, A. R. Calawa, and G. Sarmanioti, *Phys. Rev.*, 103, 247, 1956.
- Lüst, R., and J. A. Simpson, *Phys. Rev.* 108, 1563, 1957.
- McCracken, K. G., The cosmic ray flare effect, in press, 1961.
- McCracken, K. G., and R. A. R. Palmeira, *J. Geophys. Research*, 65, 2673, 1960.
- Meyer, P., E. N. Parker, and J. A. Simpson, *Phys. Rev.*, 104, 768, 1956.
- Piddington, J. H., *Phys. Rev.*, 112, 589, 1958.
- Pomerantz, M. A., S. P. Duggal, and K. Nagashima, *Phys. Rev. Letters*, 6, 123, 1961.
- Preliminary Report of Solar Activity*, High Altitude Observatory, Boulder, Colorado, TR 481, 1960.
- Quenby, J. J., and W. R. Webber, *Phil. Mag.*, 4, 90, 1959.
- Steljes, J. F., H. Carmichael, and K. G. McCracken, *J. Geophys. Research*, 66, 1363, 1961.
- Trainor, J. H., M. A. Shea, and J. A. Lockwood, *J. Geophys. Research*, 65, 3011, 1960.
- Webber, W. R., and J. J. Quenby, *Phil. Mag.*, 4, 654, 1959.
- Winckler, J. R., P. D. Bhavsar, A. J. Masley, and T. C. May, *Phys. Rev. Letters*, 6, 488, 1961.

(Manuscript received June 20, 1961;
revised July 21, 1961.)



Balloon Measurement of Solar Cosmic Rays at Fort Churchill, Canada, during July 1959

JAMES A. EARL

*School of Physics, University of Minnesota
Minneapolis 14, Minnesota*

Abstract. Balloon measurements, obtained over Fort Churchill, Canada, of solar cosmic ray intensity during July 1959 are reported. The intensities observed on July 11, 12, and 18 were many times as great as those observed simultaneously at Minneapolis. In particular, the excess Geiger counter rate at Fort Churchill was 100 ± 25 times as great as the corresponding rate observed at Minneapolis during a small influx of radiation which occurred at 0730 UT on July 11 (before the magnetic storm that began at 1620 UT July 11). Measurements of the dependence of excess rate on atmospheric depth indicate that the energy spectrum of the solar cosmic rays did not change significantly over a period of 10 days.

The purpose of this paper is to report observations made at Fort Churchill, Canada, with balloon-borne ionization chambers and Geiger counters during the solar cosmic ray events of July 1959. Since these observations were made at high geomagnetic latitudes, the results reflect the properties of the beam of solar particles; they are not affected by the geomagnetic modulation effects which are a major feature of balloon observations at lower latitudes. During the July 1959 events, high-latitude balloon measurements were also made at Fairbanks [Brown and D'Arcy, 1959], at Murmansk [Charakchian, Tulinov, and Charakchian, 1960], and at Resolute [Anderson and Enemark, 1960]. Winckler, Bhavsar, and Peterson [1961] have described in detail the sequence of solar flares, magnetic storms, and low-latitude cosmic ray effects arising from the large solar disturbances of July 1959. In brief, these events were associated with three class 3+ solar flares occurring on July 10 at 0210 UT, on July 14 at 0325 UT, and on July 16 at 2114 UT. The scope of this paper is limited to the presentation of data obtained during four flights at Fort Churchill and to the conclusions that follow from these data. A sequel will contain a more comprehensive discussion in which results obtained at various stations will be combined into an over-all picture of solar cosmic ray phenomena during the events.

Table 1 is a summary of the instrumentation, ceiling altitude, launch time, and period at ceiling for each of the four flights launched at Fort Churchill. Since the range of the telemetry

system (standard meteorological radiosonde equipment) was limited to about 200 miles, no data were obtained at geomagnetic latitudes which differed significantly from that of Fort Churchill (69°N). Pressure altitudes were measured with the usual balloon instrumentation based on the aneroid principle.

The counting instruments and the procedures for standardizing rates were identical to those developed by the Minnesota group during the IGY high-altitude monitoring program [Winckler, 1960]. In particular, the normalized single-counter rates and the normalized ionization-chamber rates, which are the data that will be presented in this paper, are directly comparable with corresponding rates obtained at Minneapolis during the same period [Winckler, Bhavsar, and Peterson, 1961]. The results from the nuclear emulsions indicated under instrumentation in Table 1 will be reported separately (Freier, to be published).

In the detailed presentation of data that follows, the normalized rates from each flight have been plotted so as to show clearly the dependence of the rates on both altitude and time. The data obtained while the balloon ascended to ceiling altitude are presented as a log-log plot of rate vs. pressure. The data obtained while the balloon floated at ceiling are presented as a plot of rate on a log scale vs. universal time. During the ascent, the logarithm of the pressure is very nearly proportional to time, so that the logarithmic pressure scale may be interpreted qualitatively as a time scale. (At

TABLE 1. Summary of Balloon Flights Made at Fort Churchill, Canada, during July 1959
(All times are UT.)

Flight No.	Date	Launch Time	Period at Ceiling	Ceiling Altitude g/cm ²	Instrumentation
478N	July 11	0543	0800 to 1610	4.0	Ion chamber Single counter Nuclear emulsions
479N	July 12-13	1049	1309 to 0343	2.5	Single counter Nuclear emulsions
481N	July 18	1152	1400 to 1640	9.7	Ion chamber Nuclear emulsions
483N	July 20	1250	1445 to 2000	4.0	Ion chamber

the top of each figure some times are given for calibration of this rough time scale.) The ascent and constant-level plots are juxtaposed so that the plot as a whole gives a qualitatively correct picture of the rate vs. time while presenting the rates during ascent as a function of atmospheric depth.

Flights 478N and 479N were aloft during the period immediately following the class 3+ solar flare that began at 0210 UT on July 10. The data from these flights are of interest because they show, first, that at times the intensity of solar cosmic rays observed at Fort Churchill was many times as great as that observed simultaneously at Minneapolis, and, second, that changes in solar cosmic ray intensity associated with the magnetic storm that began at 1623 UT on July 11 were much less pronounced at Fort Churchill than at Minneapolis.

Flight 478N reached its ceiling altitude of 4.0 g/cm² at 0802 UT on July 11 (30 hours after the July 10 flare). Figure 1 is a plot of the normalized counting rates of the single counter and of the ionization chamber. The ascent curves for flight IGC-12, launched at Minneapolis at 0700 UT July 10, are included to show the characteristics of a flight during which no solar cosmic rays were present. The excellent agreement between the rates at large atmospheric depths where only galactic cosmic rays are present shows that the data at Fort Churchill have been correctly normalized to those at Minneapolis. (Winckler [1960] has shown that, for latitudes greater than that of Minneapolis, there is essentially no latitude effect for galactic cosmic rays at the present time of maximum solar activity.) At small atmospheric depths, the large upward deviation of the curves for

flight 478N indicates that a large flux of solar cosmic rays was present. On flight 478N the single counter rates at atmospheric depths less than 10 g/cm² are high enough so that they may be affected by errors arising from saturation effects (these effects will be discussed in more detail later). Nevertheless, the fact that the rate remained practically constant during the 1-hour period during which data were obtained at ceiling indicates that the intensity was constant during this period and, by extrapolation, during the ascent. (For the relatively low rates at ceiling, spurious time variations arising from changes in the saturation effects as the counter ages are unimportant.)

Flight 479N reached its ceiling altitude of 2.5 g/cm² at 1309 UT on July 12 (59 hours after the July 10 flare). Figure 2 shows a plot of the normalized single-counter rate during the ascent and constant-level parts of the flight. It is apparent that the ascent curve for this flight is similar to that of flight 478N (and to other flights during which solar cosmic rays were present) for atmospheric depths greater than 10 g/cm². The behavior of the rates obtained at depths less than 10 g/cm² (including the rates obtained at ceiling) is dominated by saturation effects that are present when Geiger counters of the type used on these flights are exposed to high radiation intensities.

The nature of these effects is indicated in Figure 3, which shows the results of laboratory tests on a counter identical to the flight counters, which was exposed for a total of 27 hours to a radioactive source that produced an initial counting rate of 1600 counts per second. (This rate is about the same as that observed at ceiling on flight 479N.) In Figure 3 are plotted saturation

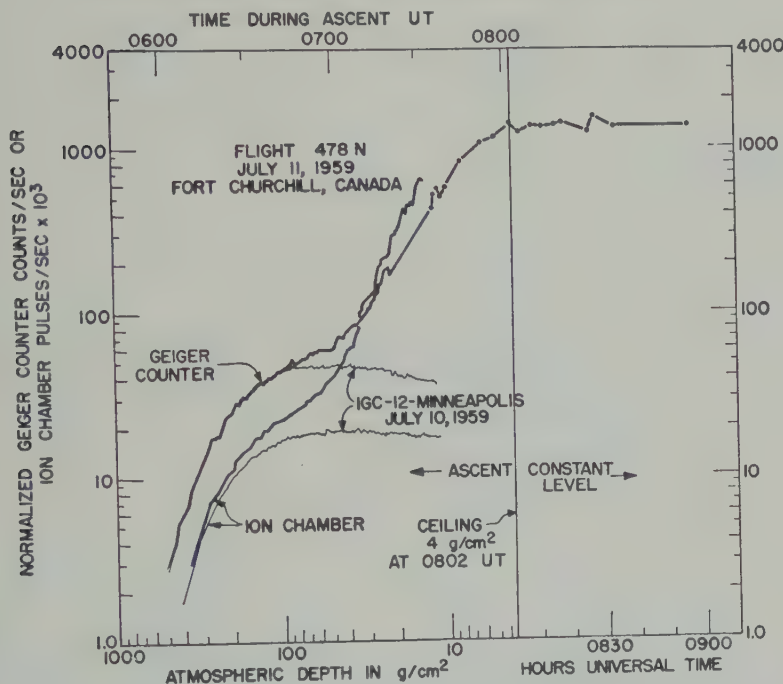


Fig. 1. Ion-chamber and Geiger-counter rates during flight 478N, which reached ceiling 30 hours after the flare at 0210 July 10. The Geiger-counter rates at altitudes above 10 g/cm^2 are subject to errors arising from saturation effects, but they show that the intensity was constant during the constant-level part of the flight. Ascent curves from flight IGC-12 show the characteristics of a normal flight.

ion curves (i.e., counting rate observed vs. true counting rate expected in the absence of saturation effects) which were measured at intervals during the exposure. It is apparent from Figure 3 that the effects are large, that they cannot be eliminated by a simple dead time correction, and that they change rapidly as the counter ages.

Although the complexity of the saturation phenomena makes it impossible to assign reliable correction factors for any rates obtained at depths less than 10 g/cm^2 , it is clear that both the decrease in the slope of the ascent curve at small depths and the rather abrupt decrease in rate that occurred late in the flight are better attributed to the phenomena presented in Figure 3 rather than to real effects. (For example, even if the true rate remained constant, the observed rate would decrease with time as the point representing the rates moved down a line such as AB in Figure 3.) Even so, the fact that there were no rapid fluctuations in the rates observed at ceiling during flight 479N means that any

variations in intensity were of a gradual nature because the characteristics of the counter are such that rapid intensity changes of a factor of 2 or more would show up as fluctuations in the observed rate. It is clear that the true counting rate was high throughout the period during which 479N was at ceiling.

A comparison of the counter rates on the ascent portions of flights 478N and 479N with the integrated fluxes measured by nuclear emulsions carried on these flights does give some information on the time variation of the intensity during 479N. The intensity as measured by counters during the ascent portion of 478N was about 1.3 times greater than that during 479N; the integrated flux as measured by nuclear emulsions during 478N was 2 times as great as that during 479N. Since data from other high-latitude flights [Brown and D'Arcy, 1959; Charakchian, Tulinov, and Charakchian, 1960] indicate that the intensity was constant throughout 478N, and, since the counter observations indicate that the intensity variation during 479N

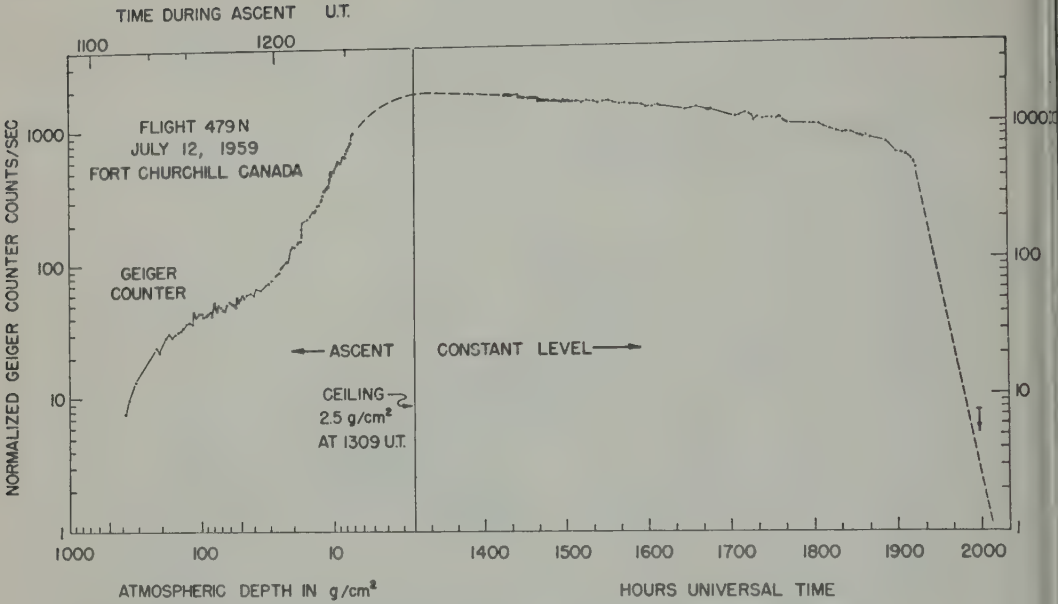


Fig. 2. Geiger-counter rates during flight 479N, which reached ceiling 59 hours after the July 10 flare. During the constant-level part of the flight, the slow decrease in rate followed by an abrupt drop at 1920 UT is attributable to aging of counter. Nevertheless, the constant-level plot shows that no rapid intensity fluctuations were present.

was gradual, it is reasonable to assume that the intensity during 479N was represented by a smooth curve which started at a value 1.3 times less than that of 478N and fell to such a value at the end of the flight that the integrated

intensity was 2 times less than that during 478N. Assuming a linear variation of intensity with time, this argument implies that the intensity dropped by a factor of 3 during the 10.5 hours that 479N was at ceiling.

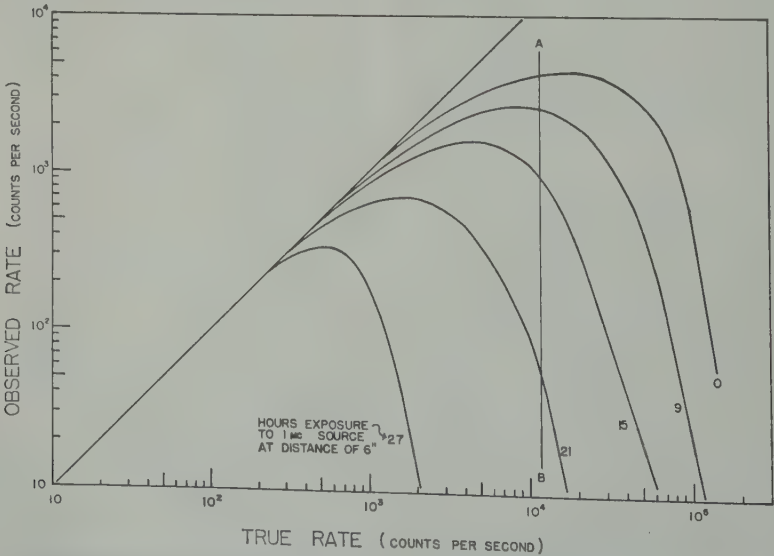


Fig. 3. Results of laboratory tests on a sample Geiger counter exposed to high radiation intensities.

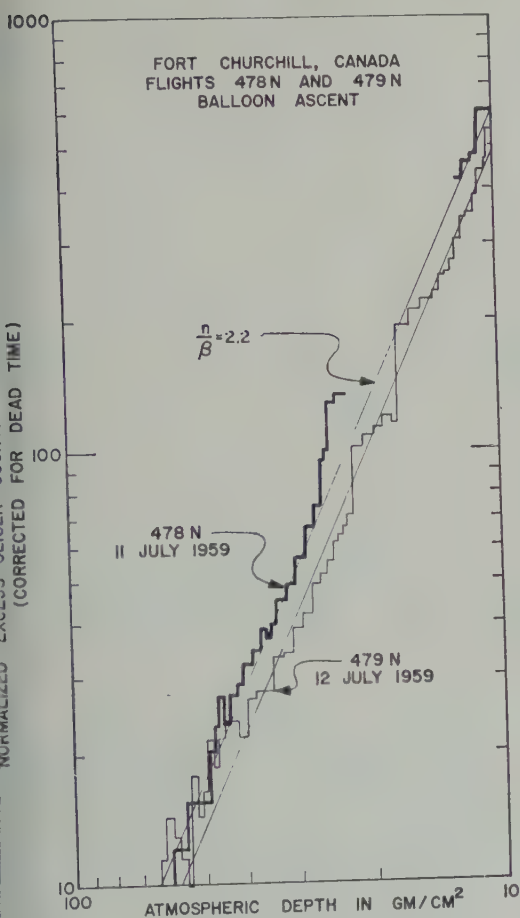


Fig. 4. Plots of excess Geiger-counter rate vs. atmospheric depth for flights 478N and 479N. Although the intensities were slightly different during the two flights, the depth dependences were nearly identical.

In Figure 4 the normalized excess rates of the single counters on flights 478N and 479N have been plotted as a function of atmospheric depth. In computing these excess rates the single counter rates on flight IGC-12, which was launched at Minneapolis on July 10 [Winckler, Bhavsar, and Peterson, 1961], were assumed to represent the contribution of normal galactic cosmic rays. This procedure is justified because the latitude effect is small and because the rate due to solar cosmic rays over most of the range of depths presented in Figure 4 is overwhelmingly large compared with the normal rate. On flight 479N the normal rate from flight IGC-12 was also corrected for the effects of a Forbush decrease present at the time of the flight. A correction for

counter dead time based on a value for the dead time of 200 microseconds was also applied in computing the rates presented in Figure 4. For low rates, this procedure is a valid method of eliminating the saturation effects discussed earlier. No correction was made for time variations during the ascent, because there is evidence that intensity changes were negligible during both ascents.

It is clear from Figure 4 that, while the rate during 478N is slightly larger than that during 479N (a factor of 1.3), the dependence of excess rate on atmospheric depth was the same in both flights and that this dependence is well represented by a power law (Rate \propto Depth $^{-u}$) with exponent $u = 2.2 \pm 0.1$. It can be shown [Winckler and Bhavsar, 1960] that, under the assumption that the solar cosmic rays are protons arriving at the top of the atmosphere with a constant isotropic flux, the counting rate of a single counter is given by Rate \propto Depth $^{-n/\beta}$, where n is the exponent in the integral energy spectrum of the protons and β is the exponent in the relationship Range \propto Energy $^\beta$, which is valid for solar proton energies if $\beta = 1.7$. If the assumptions are correct, the observed value of $u = n/\beta$ leads to the value

$$n = 3.74 \pm 0.2$$

for the exponent of the integral proton energy spectrum. The depth dependence of the excess ion-chamber rate during flight 478N is consistent with this value of n and will be discussed later in connection with two other flights carrying ion chambers.

The data obtained during flights 478N and 479N (supplemented by other high-latitude results for the initial rise in intensity) give the following rough picture of the time variations of the solar cosmic ray intensity at Fort Churchill after the July 10 flare: After an initial rise (which ended about 18 to 24 hours after the flare), during the period from 30 to 60 hours after the flare, the intensity remained practically constant (dropped less than 30 per cent) at a level such that the counting rate at 10 g/cm 2 was about 20 times normal; this period of nearly constant intensity was followed by a period during which the intensity was decreasing at the relatively rapid rate of a factor of 3 in 10 hours. Although it is not the purpose of this paper to correlate the results obtained at various stations, it is of

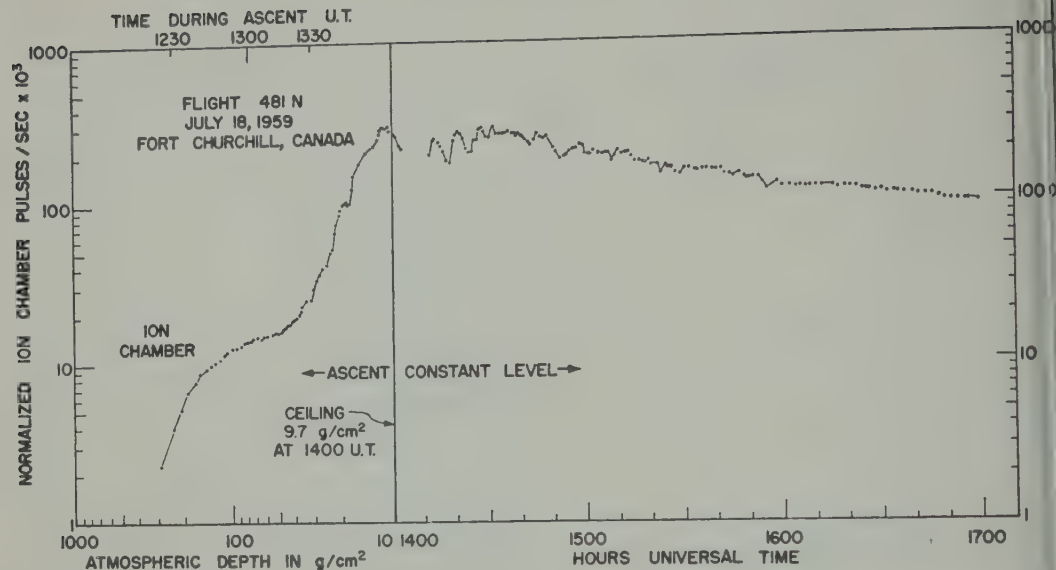


Fig. 5. Ion-chamber rates during flight 481N, which reached ceiling 41 hours after the flare at 2114 UT July 16. Note the rapid fluctuations which disappeared at the beginning of a magnetic disturbance at 1530 UT.

special interest to compare the time variations of solar cosmic rays at Minneapolis with the picture outlined by the Churchill results (see *Winckler, Bhavsar, and Peterson* [1961]). Except for a short period from 0000 to 0800 UT July 12 during which the intensities at Minneapolis were comparable to the large values observed at Fort Churchill, the intensity at Minneapolis was never more than 50 per cent above normal. While the period of high intensity at Minneapolis was most probably a result of changes in cutoff rigidity associated with the magnetic disturbance which started with a sudden commencement at 1623 UT July 11, one Minneapolis flight (IGC-13) recorded a small transient increase well before the sudden commencement. During this increase, at 0730 UT July 11, the excess single counter rate at 10 g/cm² recorded at Churchill during flight 478N was 100 ± 25 times as great as that recorded at the same depth and time at Minneapolis during flight IGC-13. This large ratio indicates that the geomagnetic field was very effective in excluding from Minneapolis the low-momentum particles that are predominantly responsible for the high intensities recorded at high latitudes. Since the rates at Fort Churchill (and at other high-latitude stations) were observed to remain nearly constant during a period while

the prestorm excess at Minneapolis decreased by a factor of 3 or more, it is apparent that this decrease is not due to a decrease in the intensity of the solar cosmic rays reaching the earth; it must be attributed either to an increase of the cutoff rigidity at Minneapolis or to a steepening of the momentum spectrum of the solar cosmic rays. The ratio of excess counter rate at Churchill to that at Minneapolis was also greater than 100 at a depth of 10 g/cm² at 1230 UT July 12 at a time while the magnetic disturbance was still present but when the high intensity observed at Minneapolis during the early part of the storm had dropped to a near-normal value (flights 479N and M-2).

Flights 481N and 483 N were launched during the period following the third class 3+ solar flare that began at 2114 UT on July 16. On both these flights, an ionization chamber was the only counting instrument used.

Flight 481N reached its ceiling altitude of 9.7 g/cm² at 1400 UT on July 18 (41 hours after the July 16 flare). Figure 5 shows a plot of the normalized ionization-chamber rate during the ascent and constant-level parts of the flight. During the ascent and the first hour at ceiling there were rapid (10-minute-period) fluctuations of about 50 per cent in the intensity, but for the

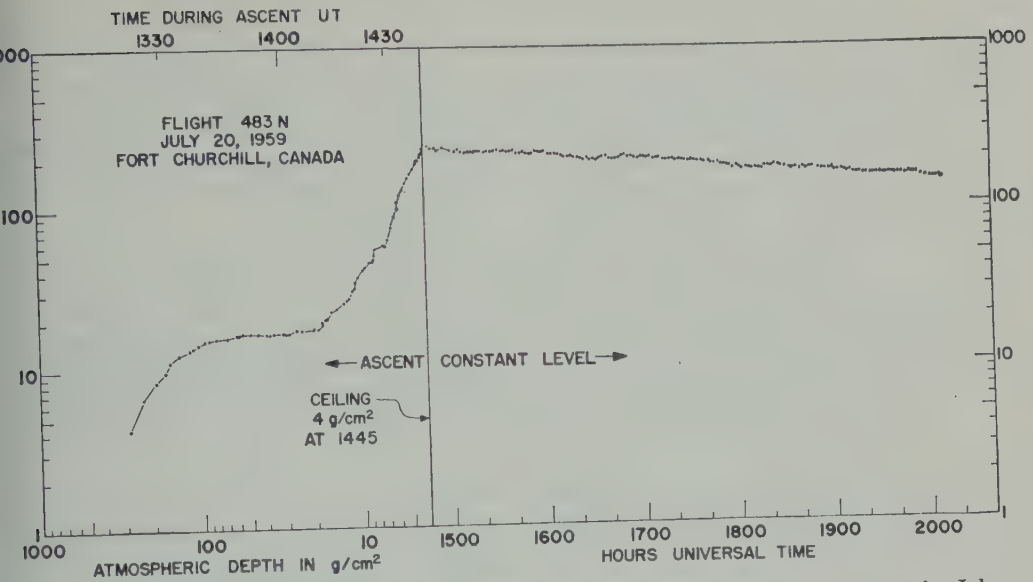


Fig. 6. Ion-chamber rates during flight 483N, which reached ceiling 89 hours after the July 16 flare. Note the extremely slow decrease of solar cosmic ray intensity during the constant-level part of the flight.

rest of the period at ceiling the intensity decreased smoothly at a rate of a factor of 2 in 2 hours. This period of rapid intensity fluctuations is unique in that it is the only time during the 8 hours that data were obtained at high altitudes from the four flights at Fort Churchill that the

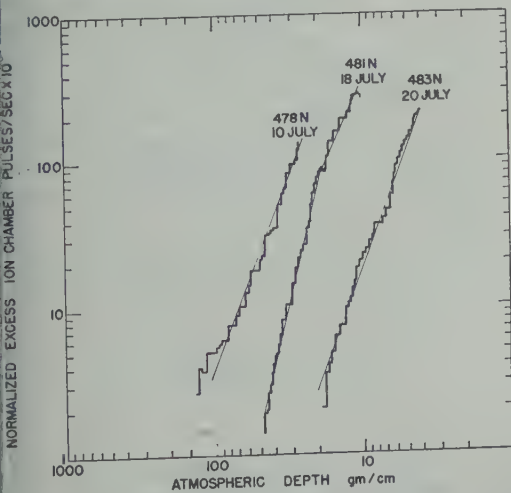


Fig. 7. Plots of excess ion-chamber rate vs. atmospheric depth for three flights at various times during July. The discontinuity in the slope of the curve for flight 481N is probably due to time variations during ascent.

intensity did not either remain constant or decrease slowly and uniformly with time. At Minneapolis, on the other hand, rapid changes arising from magnetic effects are the rule rather than the exception. It is unlikely that the fluctuations were due to auroral X rays superimposed on the solar cosmic rays, because, first, at times the counting rate was well below that expected for solar cosmic rays from a backward extrapolation of the smooth decay observed in the last part of the flight, and, second, the presence of fluctuations at atmospheric depths up to 30 g/cm² is not consistent with the rapid rate of absorption characteristic of auroral X rays. Thus it is reasonable to attribute these fluctuations to real changes in the cosmic ray intensity, but it is not clear whether these changes are present in the beam arriving at the earth or are a result of local modulation effects. In this connection, it is worth noting that the fluctuations were present just before an abrupt positive excursion of the equatorial geomagnetic field which occurred at 1530 UT and that they disappeared after this magnetic disturbance.

Flight 483N reached its ceiling altitude of 4.0 g/cm² at 1445 UT on July 20 (89 hours after the flare on July 16). Figure 6 is a plot of the normalized ionization-chamber rate during the

ascent and constant-level parts of the flight. The ascent curve shows the large increase at small atmospheric depths that is characteristic of solar cosmic rays. During the period at ceiling the intensity shows the same smooth decay observed late in flight 481N, but the rate of decay (a factor of 2 in 6 hours) was much less than that observed during 481N. It is not possible to fit the decay curves for both flights 481N and 483N by any simple time dependence (such as an e^{-kt} or t^{-n} law). Thus the decay in intensity at Fort Churchill after the July 16 flare is not described by the same t^{-n} law that Anderson found from observations at Resolute during the same period [Anderson and Enemark, 1960]. The 10 g/cm² excess intensity at Minneapolis, at the same time as flight 481N (Minneapolis flight IGC-17), was at least 50 times less than that at Churchill, although high intensities were observed at Minneapolis shortly before flight 481N was launched.

Figure 7 shows the normalized excess ionization-chamber rate as a function of atmospheric depth for the three flights on which ionization-chamber data were obtained (478N, 481N, and 483N). These curves are corrected for the effects of gradual time variations, but for flight 481N it is likely that there is a residual effect arising from the fluctuations encountered on that flight. It is apparent that the depth dependences during flights 478N (July 11) and 483N (July 20) are represented by similar power laws despite the fact that the intensities were widely different (a factor of 80). This agreement, in conjunction with the fact that the depth dependences of the excess Geiger counter rates during flights 478N and 479N were very similar, suggests that the energy spectrum of the excess radiation at Fort Churchill did not change appreciably during the 10 days that observations were made. It is apparent from Figure 7 that the excess ionization during the ascent of flight 483N was still increasing rapidly with decreasing depth even at 4.0 g/cm². This implies that the solar proton energy

spectrum increases rapidly with decreasing energy even for proton energies as low as 60 Mev.

Acknowledgments. This work was supported by the Office of Naval Research under contract Nonr-710(19). The Office of Naval Research provided balloon flights under the Skyhook program. Commander J. W. Sparkman was responsible for all the flight arrangements.

Our use of the facilities at Fort Churchill was made possible by the cooperation of Dr. R. Montalbetti of the Defence Research Northern Laboratory and of Colonel N. J. W. Smith, base commander. Special credit is due Ed Jost of Raven Industries for successfully launching and recovering the balloon flights under unusually difficult weather and terrain conditions. Early in the events, Dr. Kinsey A. Anderson, who was waiting at Fort Churchill for a flight to Resolute, provided much help and encouragement. Dr. Phyllis Freier made available before publication her emulsion results from these flights reported here. A. H. Buetow assisted in preparing the flights and in operating the telemetry receiver. The painstaking work of reading out and plotting the data was done by Roger Marth.

REFERENCES

- Anderson, K. A., and D. C. Enemark, Observations of solar cosmic rays near the north magnetic pole, *J. Geophys. Research*, **65**, 2657-2671, 1960.
- Brown, R. R., and R. G. D'Arcy, Observations of solar flare radiation at high latitude during the period July 10-17, 1959, *Phys. Rev. Letters*, **3**, 390-392, 1959.
- Charakchian, A. N., V. F. Tulinov, and T. N. Charakchian, Cosmic rays emitted by the sun, p. 649, *Space Research, Proc. First Intl. Space Sci. Symposium, Nice*, 1195 pp., North-Holland Publishing Co., Amsterdam, 1960.
- Winckler, J. R., Balloon study of high-altitude radiations during the International Geophysical Year, *J. Geophys. Research*, **65**, 1331-1359, 1960.
- Winckler, J. R., and P. D. Bhavsar, Low-energy solar cosmic rays and the geomagnetic storm of May 12, 1959, *J. Geophys. Research*, **65**, 2637-2655, 1960.
- Winckler, J. R., P. D. Bhavsar, and L. E. Peterson, The time variations of solar cosmic rays during July 1959 at Minneapolis, *J. Geophys. Research*, **66**, 995-1022, 1961.

(Manuscript received May 29, 1961;
revised July 6, 1961.)

The Calculation of the Electron Density in the Ionosphere from Elevation-Angle Measurements on Artificial Satellites

J. E. TITHERIDGE

*Seagrove Radio Research Station
University of Auckland, New Zealand*

Abstract. Measurements of the elevation angle of signals received from a satellite below the peak of the F layer enable the electron density at the height of the satellite to be calculated directly with an accuracy of a few per cent. The variation of the electron density with height can then be determined up to the peak of the F layer. Full allowance is made for the effects of refraction and of curvature of the earth. The results are largely independent of the presence of small irregularities in the ionosphere or of large-scale horizontal gradients.

Introduction. Information about the ionosphere has, in the past, been obtained almost entirely by studying the characteristics of radio waves transmitted from the ground and reflected by the ionosphere. In the last few years, however, the utilization of artificial satellites has opened up a whole new field of investigation, using signals transmitted through the ionosphere. The fluctuations in amplitude and the Doppler frequency shift of signals received from an artificial satellite have been used by several workers [e.g., *Aitchison and Weekes*, 1959; *Ross*, 1960] to determine the mean electron density below the height of the satellite. However, apart from some measurements of the time of radio rise and set used by *Alpert, Dobriakova, Chudenko, and Shapiro* [1958] for an approximate calculation of the electron density in the upper ionosphere, very little consideration has been given to the angle of arrival of the signals.

For a satellite moving horizontally in a static, spherically symmetrical ionosphere it can be shown that the Doppler frequency shift Δf of the received signal is given by $\Delta f/f = (U/c) \cos \Phi \cos \delta$, where U is the velocity of the subsatellite point, Φ is the angle between the direction of movement of this point and its bearing from the receiver, and δ is the elevation angle of the received signal. In this expression, only the term $\cos \delta$ depends on the ionosphere. Measurements of the elevation angle of the received signal therefore give exactly the same ionospheric information as the more usual Doppler measurements. Measurement of $\cos \delta$ has the advantage, however, that an absolute value is obtained at each instant, so that useful

measurements can be made over considerable distances. The value of $\cos \delta$ is also independent of any vertical motion of the satellite and of changes in the ionosphere. The use of elevation-angle measurements also removes the need for the accurate calculation of the rapidly changing term $\cos \Phi$, so that the orbit of the satellite need not be known as precisely as in the analysis of Doppler measurements.

In the present work the value of $\cos \delta$ is determined directly, with an accuracy of 1 per cent, by means of a rotating interferometer of the type described by *Whale* [1954]. The interferometer has two aerials mounted on the ends of a 60-foot horizontal arm which rotates once every 10 seconds and produces a continuous record of the bearing and elevation angle of the received signal. The measured values of $\cos \delta$ could be analyzed by the methods employed by other workers for the analysis of Doppler measurements. However, these methods make assumptions that restrict the calculations to near-vertical incidence, so that only the mean electron density below the height of the satellite can be determined. Corrections are also required for the effects of refraction and of curvature of the earth. In this paper a method of determining the actual path followed by the radio waves in the ionosphere is employed. It makes no assumptions about the shape of the ionized layers, but gives directly the electron density at different heights in the ionosphere.

This paper will consider only records taken during the last month in the life of the satellite 1958 δ_2 (Sputnik III). During this period the satellite passed over New Zealand at a height

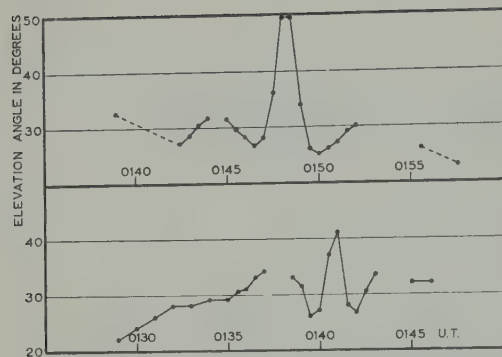


Fig. 1. The angle of arrival of the 20 Mc/s signals from 1958 δ_2 , March 29 and March 30, 1960.

of about 250 km, well below the peak of the F layer. In this case the elevation angles measured when the satellite is nearly overhead depend only on the mean electron density below the height of the satellite. At greater ranges the measured elevation angles depend more on the electron density near the satellite, while the lowest observed angle depends only on the electron density at the height of the satellite. The electron density at this height can therefore be determined directly and taken as a convenient starting point for the analysis.

The effect of the magnetic field will be neglected throughout this paper. It can be approximately allowed for by multiplying the calculated values of the electron density by $1 \pm (f_H/f) \cos \alpha$, where f_H is the gyrofrequency, α the angle between the direction of the magnetic field and the horizontal direction of propagation, and the $+$ and $-$ signs refer to the ordinary and extraordinary rays, respectively. For a frequency f of 20 Mc/s and a magnetic dip angle of 60° this correction amounts to less than 3 per cent for the ordinary ray. Since the measured elevation angles give the average of the values for the ordinary and extraordinary rays, the final correction for the effect of the magnetic field should seldom exceed 1 per cent.

The electron density at the height of the satellite. The elevation angles measured during two typical daytime transits are shown in Figure 1. This figure shows clearly the two features that were always observed during near-overhead transits: the disappearance of the signal about 4 minutes before and after the time of closest approach, and the occurrence of a well-defined

minimum elevation angle at a ground range of about 600 km.

Both these effects are readily explained by considering the possible ray paths between a satellite below the peak of the ionosphere and a receiver on the ground. Three such paths are shown in Figure 2, calculated for a parabolic F layer typical of that existing over Auckland during March. The heavy line gives the highest path that is likely to give a useful signal; signals received by any higher path would be greatly attenuated because of the small electron gradient at reflection, giving increased absorption and a large virtual height of reflection. The light line gives the path that just reaches the height of the satellite; the broken line gives the path with the smallest ground range.

Figure 2 shows that signals will be received directly from the satellite up to a ground range of about 1200 km, but that reception is not possible at ranges between 1200 and 1600 km. At greater ranges the signals are received after one or more ground reflections. At ranges less than 1200 km the elevation angle of the received ray decreases to a minimum value, before increasing rapidly as the satellite passes overhead. This minimum angle δ_0 corresponds to the ray that is horizontal at the height h of the satellite. Bouguer's law therefore gives $(R + h)/\mu = R \cos \delta_0$, where R is the radius of the earth and μ is the phase refractive index at the height of the satellite. The value of the plasma frequency f_N at the height of the satellite is therefore given by

$$f_N^2/f^2 = 1 - \mu^2 \\ = 1 - \cos^2 \delta_0 (1 + h/R)^{-2} \quad (1)$$

The corresponding value of the electron density is

$$N = 12400 f_N^2 \text{ electrons/cm}^3$$

when f_N is in megacycles per second.

Equation 1 assumes that the satellite is moving horizontally. If the orbit is inclined at an angle α to the horizontal, the value of $\cos \delta_0$ in equation 1 must be replaced by $\cos \delta_0 \sec \alpha$. For a near-circular orbit, however, this correction can be ignored, since for α less than 5° it corresponds to an error of less than 0.4 per cent in $\cos \delta_0$. The effect of horizontal gradients of ionization can be removed, to a first approximation, by averaging the two values of δ_0 observed in each

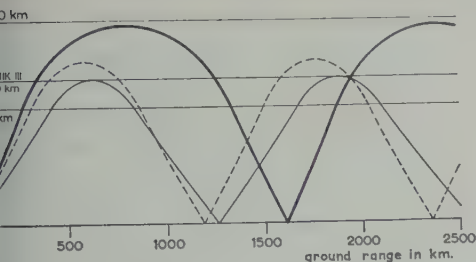


Fig. 2. Ray paths for a curved earth and a dispersive F layer. $f = 20$ Mc/s; $f_c = 13$ Mc/s. Solid curve, the highest useful ray ($\delta = 36^\circ$); dashed curve, the lowest possible ray ($\delta = 24^\circ$); dotted curve, ray with the smallest ground range ($\delta = 31^\circ$).

height. Equation 1 then gives the value of f_N at the point of closest approach of the satellite, provided that f_N varies approximately linearly with height over the interval (of about 10 km in the present work) through which the satellite passes between the two observations of δ_0 .

The measurements shown in Figure 1 give $f_N = 10.0$ Mc/s at a height of 243 km on March 29 and $f_N = 9.9$ Mc/s at 238 km on March 30. The mean values of $\cos \delta_0$ can be obtained to within 1 per cent, so that the values of f_N have a maximum experimental error of 3 per cent. The presence of nonlinear gradients of ionization could, however, cause the over-all error to exceed this value. The results obtained do not depend on any assumptions about the form of the lower ionosphere and could therefore be used to provide a valuable check on the electron-density distributions calculated from $h'(f)$ records.

The variation of the electron density with height. The path followed by a ray in the ionosphere can be approximately determined for many commonly assumed models of the ionosphere, but for accurate work it seems preferable to use a model for which a simple, accurate solution of the ray-path equation can be found. In the present work the ionosphere is therefore supposed to consist of a number of 'near-linear' sections, in each of which the value of $\mu^2 r^2$ varies nearly with the distance r from the center of the earth. Since the fractional variation of r will be small in any one section, these sections differ only slightly from the usual linear sections in which μ^2 varies linearly with r .

Consider a ray with an elevation angle δ at the surface of the earth, of radius R . The angle i

between the ray path and the vertical at any point is then given by

$$\mu r \sin i = R \cos \delta \quad (2)$$

for a spherically stratified ionosphere. If the value of $\mu^2 r^2$ increases linearly with height from the value $\mu_1^2 r_1^2$ at $r = r_1$ to the value $R^2 \cos^2 \delta$ required for reflection at $r = r_m$, we have

$$\mu^2 r^2 = \frac{r_m - r}{r_m - r_1} \mu_1^2 r_1^2 + \frac{r - r_1}{r_m - r_1} R^2 \cos^2 \delta \quad (3)$$

Eliminating μr between equations 2 and 3, and writing

$$\mu_1^2 r_1^2 / R^2 \cos^2 \delta = 1 + K^2$$

we get

$$K \tan i = \{(r_m - r_1)/(r_m - r)\}^{1/2}$$

The ray-path equation $r d\theta/dr = \tan i$ therefore gives the angular range θ (measured at the center of the earth) between the heights r_1 and r_m as

$$\begin{aligned} \theta &= \frac{\sqrt{r_m - r_1}}{K} \int_{r_1}^{r_m} \frac{dr}{r \sqrt{r_m - r}} \\ &= \frac{2\sqrt{r_m - r_1}}{K \sqrt{r_m}} \tanh^{-1} \sqrt{1 - \frac{r_1}{r_m}} \end{aligned}$$

or

$$\theta \div (2/K)(r_m - r_1)/r_m$$

with a maximum error of 0.1 per cent for $r_m - r_1 < 350$ km. The ground range covered by a ray between the height r_1 and the maximum height reached by the ray is therefore

$$R\theta = (2R/K)(r_m - r_1)/r_m \quad (4)$$

where

$$K = \left\{ \frac{r_1^2}{R^2 \cos^2 \delta} \left(1 - \frac{f_1^2}{f^2} \right) - 1 \right\}^{1/2}$$

and f_1 is the plasma frequency at the height r_1 .

The use of this relation to determine the variation of the electron density with height will be illustrated by an analysis of the measurements made on March 29, 1960, and shown in Figure 1. The elevation angles measured at equal ranges before and after the time of closest approach are first averaged and plotted as a function of the ground range. This produces the

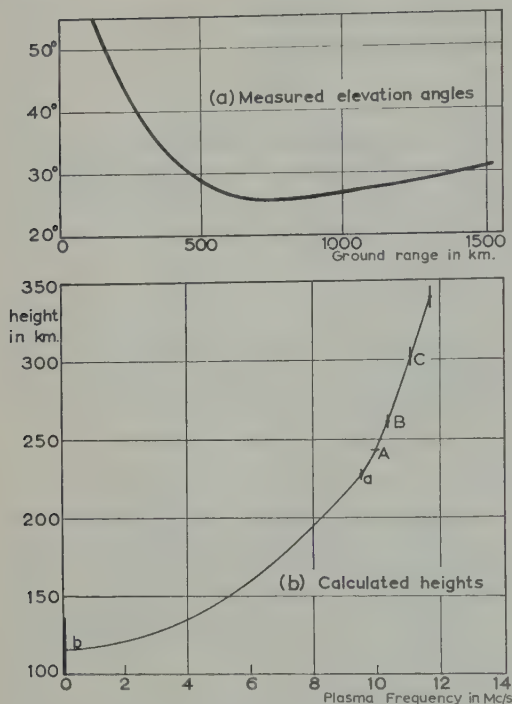


Fig. 3. The calculation of the electron density in the *F* layer. March 29, 1960, 0150 UT.

smooth mean curve shown in Figure 3(a), in which the effects of horizontal gradients and of changes in the height of the satellite have been largely eliminated. The minimum value of δ (25.6°) is then used to calculate the plasma frequency at the height of the satellite from equation 1. This gives the point *A* ($f_N = 9.95$ Mc/s, $h_A = 243$ km) in Figure 3(b).

The next step is to determine the maximum height reached by a ray with an elevation angle one or two degrees greater than the minimum value, say 27° . Figure 3(a) shows that this angle occurred at ranges of 504 and 954 km. A ray transmitted from the receiving site with an elevation angle of 27° would therefore reach the height of the satellite at a range of 504 km, and return to this height at a range of 954 km. The ground range covered by the ray between the height of the satellite (243 km) and the height of reflection is therefore $\frac{1}{2}(954 - 504) = 225$ km. Inserting this value of $R\theta$ in equation 4 (with $r_1 = R + 243$ km, $f_1 = 9.95$ Mc/s, and $\delta = 27^\circ$) gives $r_m - r_1 = 17$ km. The maximum height reached by the ray is therefore $r_m =$

$R + 260$ km. The plasma frequency required for reflection at this height is calculated from

$$f_N^2 = f^2(1 - R^2 \cos^2 \delta / r_m^2) \quad (1)$$

This gives the point *B* ($f_N = 10.37$ Mc/s, $h = 260$ km) in Figure 3(b).

The next point *C*, corresponding to an elevation angle of, say, 29° , is determined similarly. The ground range occurring between the height of the satellite and the height of reflection (at $\delta = 29^\circ$) is obtained from Figure 3(a). The range occurring in the section *AB* is approximately determined (assuming a linear increase of f_N^2 from *A* to *B* and neglecting the curvature of the ionosphere) and subtracted, leaving the range *Rθ* in the section *BC*. The height of this point *C* is then calculated from equation 4, and the plasma frequency at this height is obtained from equation 5.

This process is repeated to determine further points until the peak of the *F* layer is reached. If tables are prepared giving the value of $K/2.2$ as a function of f_N for three or four typical values of r_1 , at each value of δ used in the analysis, the complete calculation can be carried out in a few minutes. The results are accurate to within a few kilometers, the maximum probable error at each point being indicated by the short lines in Figure 3(b). The calculated heights are quite independent of any assumptions about the distribution of ionization below the height of the satellite.

The electron-density profile below the height of the satellite cannot be determined in detail since only the total electron content below the satellite, and the electron gradient at the satellite, are important in determining the amount of refraction. The ionosphere below the satellite is therefore assumed to consist of only two near-linear sections, from the point *b* in Figure 3(b) (where $f_N = 0$) to the point *a* ($f_N = 0.95 f_A$), and from the point *a* to the known point *A* ($f_N = f_A$). The sizes of the two steps $h_A - h_a$ and $h_a - h_b$ are calculated to give the correct range for points on the left-hand section of the curve in Figure 3(a). The calculated step from *a* to *A* gives the correct gradient of ionization at *A*, but the lower step *ab* serves only to show the total amount of ionization below the point *a*.

Acknowledgments. This work was carried out at the Seagrove Radio Research Station, a section of

Physics Department of the University of land. The station receives grants from the Zealand Department of Scientific and Industrial Research, the Post and Telegraph Department, the Broadcasting Service. The work described in this paper was supported by Research Grant 54-60 from the National Aeronautics and Space Administration, Washington, D. C.

REFERENCES

W. H. R. H. J., and K. Weekes, Some deductions of ionospheric information from the observations of satellite emissions from 1957 α_2 , *J. Atmospheric and Terrestrial Phys.*, **14**, 236-243, 1959.

W. H. R. J., I. A., F. F. Dobriakova, E. F. Chudsenko,

and B. S. Shapiro, On the results of determining the electron concentration of the outer regions of the ionosphere by means of observations of the radio signals of the first satellite, *Doklady Acad. Nauk SSSR*, **120**, 743-746, 1958.

Ross, W. J., The determination of ionospheric electron content from satellite Doppler measurements, *J. Geophys. Research*, **65**, 2601-2606, 1960.

Whale, H. A., A rotating interferometer for the measurement of the directions of arrival of short radio waves, *Proc. Phys. Soc. London, B*, **67**, 553-562, 1954.

(Manuscript received May 19, 1961;
revised July 20, 1961.)

An HF Radar Search for Possible Effects of Earth Satellites upon the Upper Atmosphere

T. A. CROFT AND O. G. VILLARD, JR.

*Radioscience Laboratory, Stanford University
Stanford, California*

Abstract. Several investigators have interpreted experimental data as evidence that the passage of artificial earth satellites creates large-scale disturbances in the ionosphere. In an effort to confirm this hypothesis, a search for such effects in the immediate vicinity of Sputnik III and Echo has been conducted with the aid of an HF radar having a comparatively broad beamwidth, both in azimuth and in elevation. Possible satellite-associated disturbances were sought both as direct reflections from the vicinity of the vehicle and as perturbations within the F layer sufficient to alter the structure and appearance of ground backscatter mirrored by that layer. In an appreciable fraction of the 139 observations made, the equipment would have registered either type of disturbance, owing to direct illumination of the passing satellite by high-angle radiation from the transmitter, and illumination of the distant ground by lower-angle energy reflected down from the ionosphere beneath the satellite's track. In all this work, the fundamental assumption was made that, if an effect were produced, it would be closely associated with the satellite passage in time, and it would also occur comparatively close to the satellite's track in space. No observed returns could be attributed to direct reflection. Many layer anomalies were found to occur at locations below orbiting vehicles, at times close to the time of vehicle passage. Subsequent study of these anomalies, and comparison with statistical and other characteristics of natural changes, did not provide any basis for believing any of the anomalies to be satellite-caused. It was concluded that all phenomena detected were of natural origin.

INTRODUCTION

Since the advent of artificial earth satellites, radar investigators have obtained experimental results which has been interpreted as evidence that orbiting vehicles might produce relatively large changes in electron-density distribution in the upper atmosphere [e.g., Kraus and Higgy, 1959].

In mid-1959 a search for such effects was undertaken, using existing HF backscatter radar equipment. This technique was selected because it permits detection of either a direct reflection from a target or a refractive effect due to ionization changes which are not dense enough to produce direct reflections (e.g., a ripple, or perturbation in the layer).

To establish a correlation between observed returns and satellite passages by means of a statistical study of many samples, 139 records were obtained during a 14-month period. They have been analyzed both individually and collectively, and the detailed results were published in *Stanford Radioscience Laboratory Technical Report 24* of March 10, 1961, under contract number 225 (33). The major portions of this report have been summarized here.

EQUIPMENT AND OPERATIONAL TECHNIQUE

All equipment used in the experiment was located near Stanford, California, at 37.4°N , 122.2°W . Two HF radar sets and four antennas were operated in various combinations. One-millisecond pulses in the 10- to 30-Mc/s frequency interval were generated 10 times per second at peak powers of 20 or 50 kw. Receiver bandwidths were typically 3 kc/s. The higher power radar was phase coherent. Of the antennas, two were conventional rhombics, one was a broadside array of 8 rhombics, and one was a steerable log periodic. Gain ranged from 8 db for the log periodic to 30 db for the rhombic array.

Satellites used in the experiment were Sputnik III (195882) and Echo I (196011). During passages of the Sputnik, 133 records were taken. This satellite was chosen because of its size, changing height, and the good orbital data available. The launch of the Echo satellite just before the end of the experiment permitted the recording of 6 passages of this large vehicle.

For each record, the predicted satellite passage geometry was used to select an antenna and its heading. Approximately 40 minutes before

passage, the radar was run at various frequencies to determine the optimum frequency for detection of effects at the ranges of interest.

If possible, the antenna and frequency were selected so that the rays would illuminate the satellite and also refract off the ionosphere near or underneath it. When satellites passed at heights above the F layer, direct illumination could not be achieved outside of a conical volume of revolution bounded by the critical angle of elevation (rays of lower elevation angle are refracted downward). Refraction from the neighboring ionosphere could always be accomplished when the satellite range exceeded half the skip distance.

Twenty minutes before satellite passage, radar operation commenced, and returning signal amplitude and phase were recorded on magnetic tape for approximately 40 minutes. This information was later transferred to 35-mm strip film to produce range-time presentations. Those records showing phase were termed 'Doppler records' since they provided a measure of the rate of change of phase path length to any given source of scatter.

METHOD OF DATA REDUCTION

After the collection and reduction of data on the satellite passage, a plan view of the event was drawn showing satellite positions and times relative to antenna positions and headings. Sputnik III position data were extracted from the ephemeris published by the Smithsonian Institution. Similar data for Echo I was acquired by using local optical sightings in conjunction with orbital predictions supplied by NASA and Space Track Control Center.

It was assumed that effects caused by the satellite would originate near it and would become visible within a few minutes of its passage. Returns existing before the entry of the satellite into the antenna pattern served as a basis for judging the normal state of upper-atmospheric ionization. (Degree of stability and number of natural anomalies differ greatly from record to record.) Anomalies in the returning signal immediately before, during, or after satellite passage were compared in range and time against the satellite range, time, and position in the antenna beam. Weighing all known factors, the film reader assigned to each record a rating number of 0, 1, 2, or 3, which

was a subjective measure of the probability that anomalies on the record were caused by the satellite.

After completion of rating assignments, various experimental parameters of possible significance were selected and the ratings of the records were compared against them. Some direct plots of individual records against the parameters were made, but owing to the large number of records, numerical methods were usually necessary to detect statistical trends.

In reading filmed backscatter records, it is often difficult to decide what constitutes an anomaly and what represents normal background behavior. Also, when an anomaly is identified, the time of its commencement is frequently difficult to determine, because of rather slow onset or extreme faintness. Consequently, it was evident that the reader knowing the position and time of the satellite might have inserted an artificial correlation. Therefore, a new reader examined all records in detail, noting the nature of anomalies and their starting times, but without access to any information about the position of the satellite. Results thus far obtained have indicated that the great majority of correlated anomalies occurred in the range from approximately 1000 to 4000 km, and so the second reader restricted his attention to anomalies in this range interval. Anomalies were divided into five classes.

Amplitude records. New returns, or the intensification of existing returns, were combined under the classification 'new returns.' A temporary decrease in the amplitude of an existing return, or a temporary diffuseness of an otherwise well-defined return, was called a 'diffuse anomaly.'

Doppler records. On these films there are often range intervals in which near-vertical striations appear. This means that in the range interval under consideration the phase path length of the dominant propagating mode is changing at a rate that corresponds in half-wave lengths per second to the number of striations lines per second. Therefore, Doppler anomalies containing striations of finer detail than have existed previously in that range interval were classified as 'fast,' and those containing coarse striations were classified as 'slow.' Some Doppler anomalies that do not fall clearly into either class can be detected; they were called 'uncertain.'

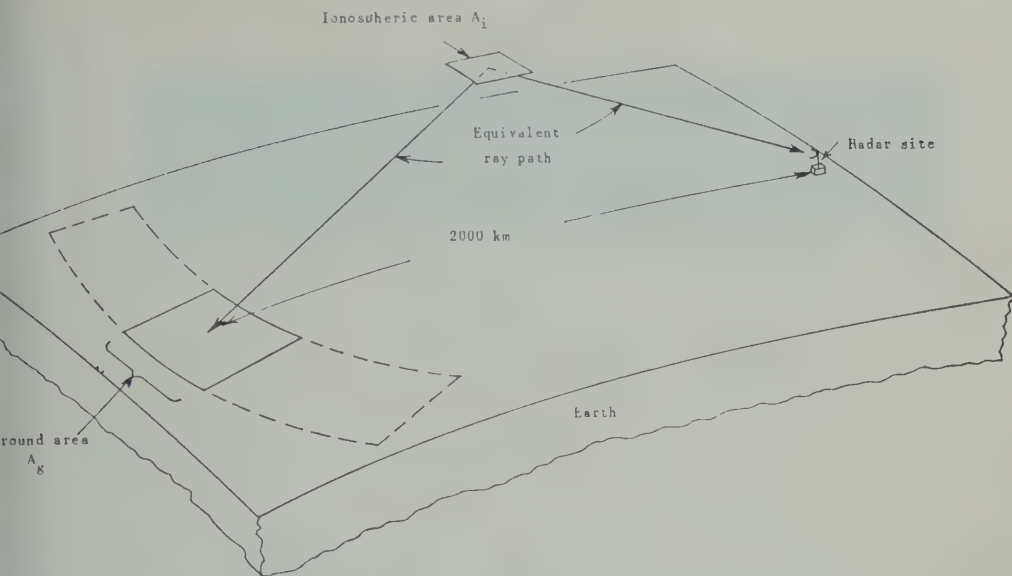


Fig. 1. Relationship between ground and ionosphere areas during backscatter.

After all films had been examined using the above criteria, satellite orbital data were made available and the time scale of each record was analyzed to T_0 , the time of the satellite's closest approach. The resulting ' T_0 time' was divided into 1-minute periods, for each of which the average number of returns per record was computed.

THRESHOLD OF DETECTION

Since calculations of the size of satellite-induced anomalies required to produce detectable backscatter effects are dependent on many assumptions and estimates, the results given in the following simplified discussion should be interpreted as an indication of order of magnitude.

Direct reflection. Assume that a satellite is 500 km from the log periodic antenna at the minimum azimuth and elevation and produces in its vicinity an ionized region which is to be detected by direct reflection at 12 Mc/s. Parameters are as follows:

- = antenna power gain over isotropic radiator = 6.5.
- = peak pulse power radiated = 5×10^4 watts.
- = distance to satellite = 10^6 meters.
- = wavelength = 25 meters.
- = Boltzmann constant.

B = receiver bandwidth = 3 kc/s.

T = temperature = 300°K .

σ = effective area of target; this is the value sought.

In the frequency interval of interest, the main sources of noise are atmospheric and cosmic. The resulting equivalent noise temperature is highly variable, having maxima and minima several orders of magnitude apart. After examination of data in the literature [Cottony and Johler, 1952; Van der Ziel, 1954; Pawsey and Bracewell, 1955; International Telephone and Telegraph Corp., 1956], an equivalent temperature of 500 T was selected for 12 Mc/s, and it was assumed that the factor varies as $\lambda^{2.3}$ for other frequencies.

Experience in film reading has led to an estimate that a return will be visible if it has a signal-to-noise ratio of 0.6 for 30 seconds (300 pulses). Neglecting absorption,

$$\sigma = 0.6 \frac{(500kTB)64\pi^3 R^4}{P_r G_a^2 \lambda^2} = 5500 \text{ m}^2$$

Notice the great effect of range; at $R = 500$ km, σ is reduced to 350 m^2 .

Six records were obtained when the Sputnik III vehicle passed through the rhombic array's main lobe at distances from 925 to 1430 km and heights under 205 km. This antenna has

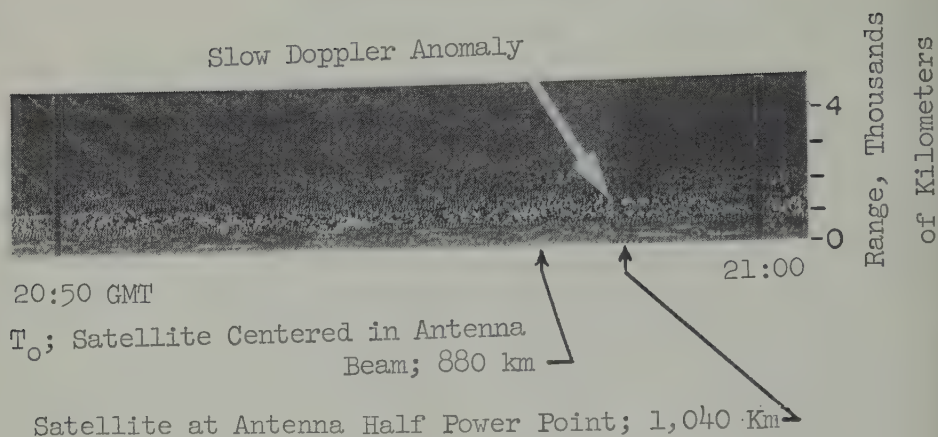


Fig. 2. Sample Doppler record.

$G_a = 500$, but the passages were under the main lobe elevation, where G_a is about 250. Typical values of radar cross section required for detection under these conditions are in the range 10 to 100 m² (e.g., $\sigma = 13$ m² for $P_r = 2 \times 10^4$, $\lambda = 19$ meters, and $R = 1100$ km). Despite this high sensitivity, no anomalies were observed to have the same range and radial velocity as the satellites.

Ionospheric anomalies. Most of the energy received by this backscatter technique is being returned by fixed land surfaces; thus changes observed are caused by changing ray paths, which in turn are due to variations in electron density and gradient. Using the simplified assumptions illustrated in Figure 1, assume that electron density in ionospheric area A_i is initially too low to reflect the radar beam to ground

area A_g . If, through some mechanism, the electron density is increased sufficiently to cause the indicated reflection, the order of magnitude of A_i required to support a detectable echo can be computed. It is assumed that the ground scatters 1 per cent of the energy incident upon it. With the experimental conditions used in the first radar cross-section computation, the minimum ground area must be roughly 40 km². The corresponding ionospheric area, A_i , must then be 10 km². Use of the higher-gain rhombic reduces this area considerably. The magnitude of the electron-density increase required depends on the initial conditions. At the height at which refraction is taking place, a 1 per cent increase represents an addition of approximately 10^{10} electrons per cubic meter.

To produce a detectable anomaly in existing

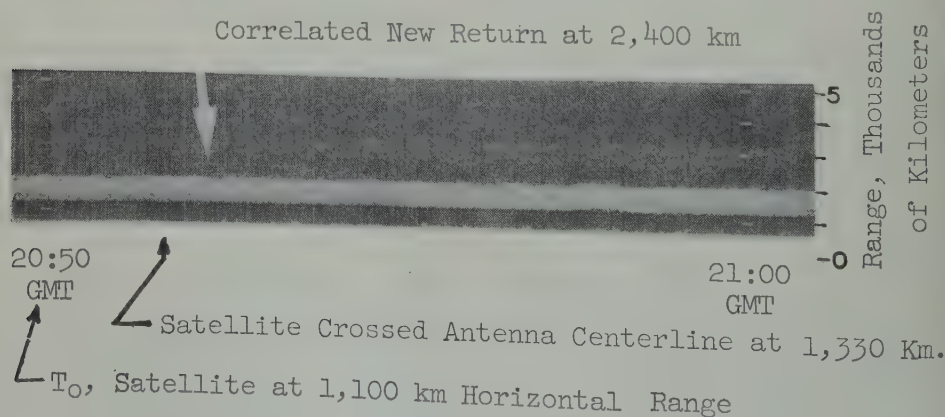


Fig. 3. Sample amplitude record.

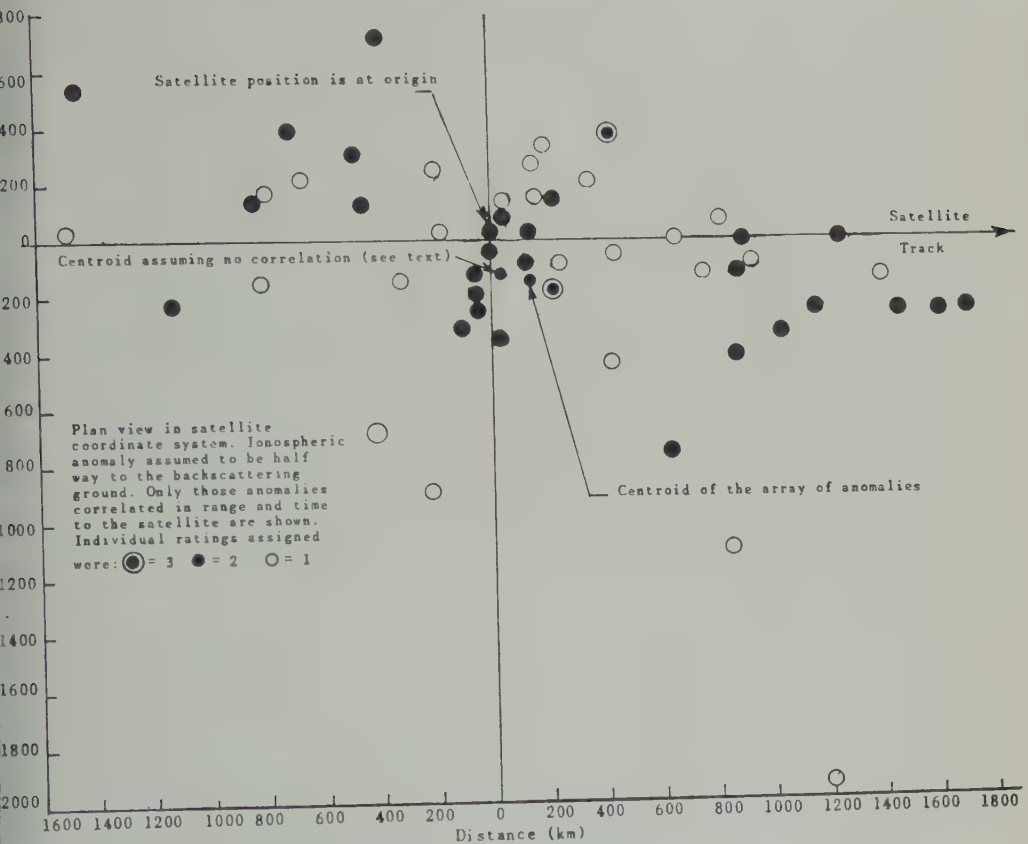


Fig. 4. Location of correlated ionospheric anomalies relative to the satellite.

backscatter, the refracting characteristics of the ionosphere must be changed over a much greater area than that discussed above, since backscatter from undisturbed regions constitutes the 'noise' against which the changed signal must be detected. If the entire ionosphere is assumed to be reflecting radar waves, then, at some time during a pulse interval, the radar is illuminating an area shown within the dotted outline on Figure 1. The 1-msec pulse is approximately 300 km long. The resulting ground area illuminated is then approximately $10,000 \text{ km}^2$ per degree of antenna beamwidth at a distance of 2000 km, and is proportional to that range. Although this is a large area, several mechanisms operate to reduce the size of a minimum detectable variation.

Smaller density fluctuations would probably be required to change the refracting characteristics of a layer than would be required to

cause a nonreflecting layer to become reflecting.

In many range intervals, it is apparent from the backscatter records that only a part of the ionosphere is acting as the idealized equivalent reflecting sheet. This reduces the background return against which an anomalous return must be detected.

The Doppler technique permits the detection of changes in the relative magnitude of various propagating modes. Thus if two approximately equal reflections are being obtained from different ionospheric areas and some effect causes one of them to be enhanced, the phase characteristics of the total received signal may change in a detectable manner. Often in data reduction it was found that an anomaly visible on a Doppler record in an existing backscatter return was not seen in the corresponding amplitude record.

Actual rays return to earth by refraction rather than by the reflection illustrated in

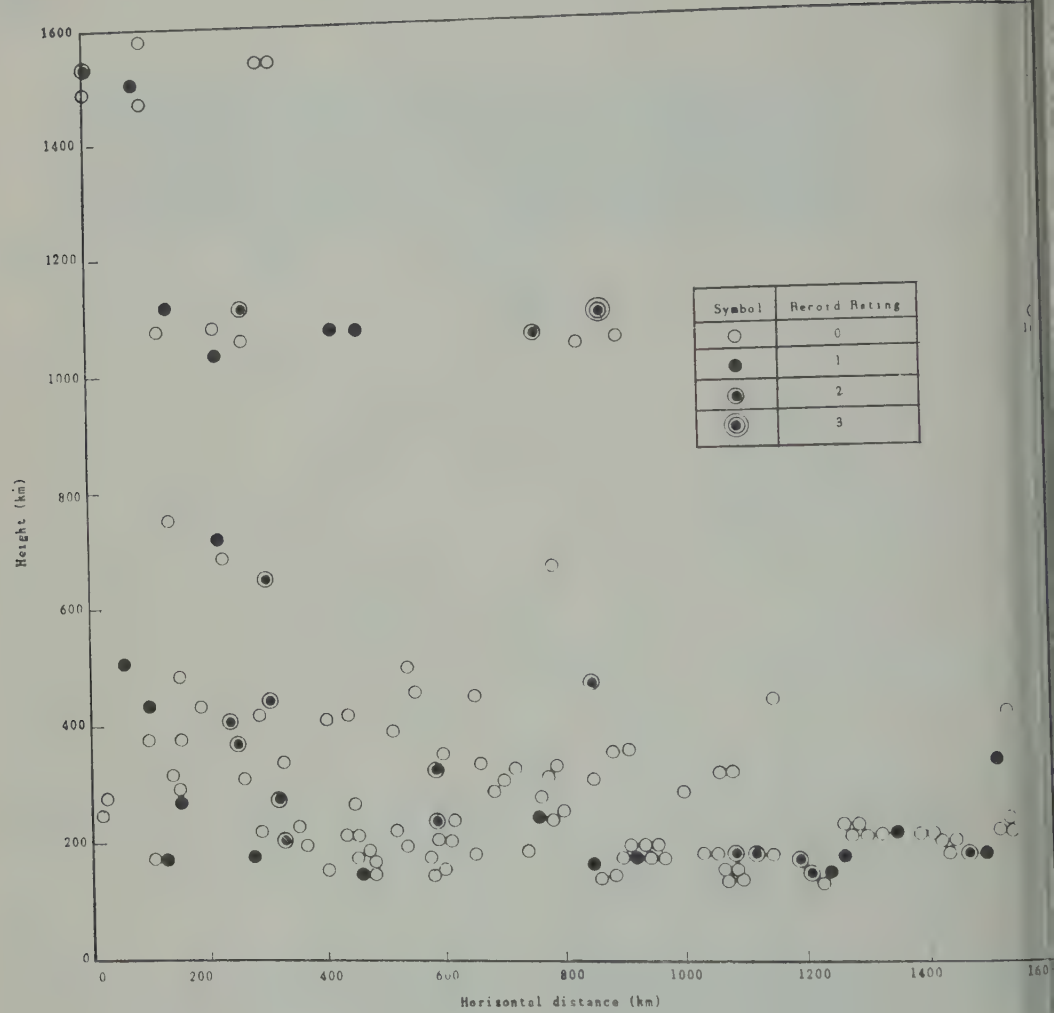


Fig. 5. Record ratings vs. height and distance of nearest approach of satellite.

Figure 1. In doing this, they undergo a focusing process near their apex due to the greater bending of the higher-elevation rays at a fixed distance from the radar. Thus a relatively small variation in electron-density distribution in the upper reaches of the ray paths would affect rays which subsequently strike the ground over a large area.

RESULTS OF DATA REDUCTION

No returns were detected which were attributed to direct reflection from the vicinity of the satellites. However, many changes in ground backscatter were found at ranges and times such that the corresponding ray paths passed under

or in the vicinity of the satellite. On the basis of these, ratings were assigned in the following numbers:

Rating	0	1	2
Number of records	101	19	18

Some of the anomalies appearing on the higher rated records correlated quite well in range and time to the passage of the satellite, and these taken by themselves, could have been used to build a fairly impressive case for the presence of satellite-caused ionospheric perturbations. Typical of the records are shown as Figures 2 and 3.

The Doppler film shown was rated '3' and was obtained by using the log periodic antenna

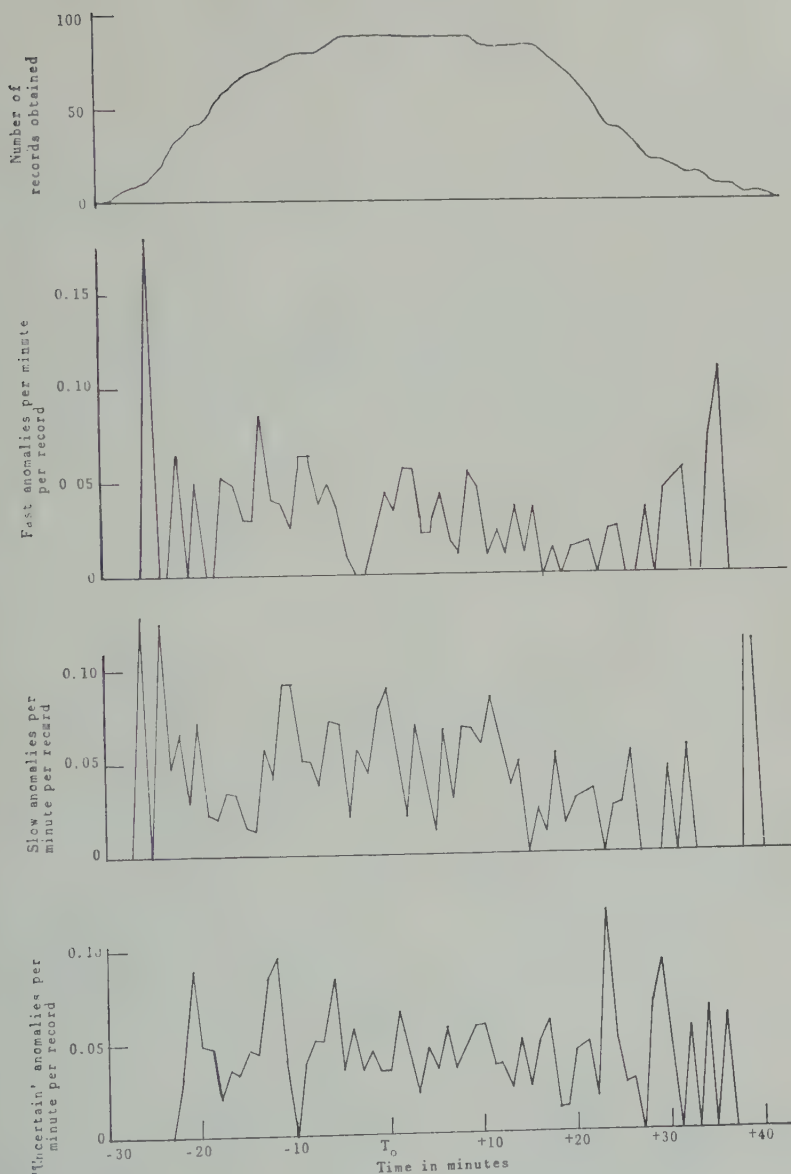


Fig. 6. Doppler anomalies vs. T_0 time.

ing directly at the satellite's point of nearest approach. A clear, unique slow Doppler is seen very nearly twice the horizontal range of the ellipse.

The amplitude record was taken with the omnibic array directed 30° down range from the ellipse's nearest approach. The clear new return shown was followed by another, weaker return at the same range 15 minutes later; therefore the film was rated only '2.'

Many changes occur on backscatter records such as these at times and ranges that do not correlate with satellite passage. Therefore, the collected group of correlated returns were further examined to see whether they exhibited some statistical characteristic that would indicate whether or not they had been caused by satellite passages.

Figure 4 shows the location of all recorded satellite-correlated anomalies plotted on the

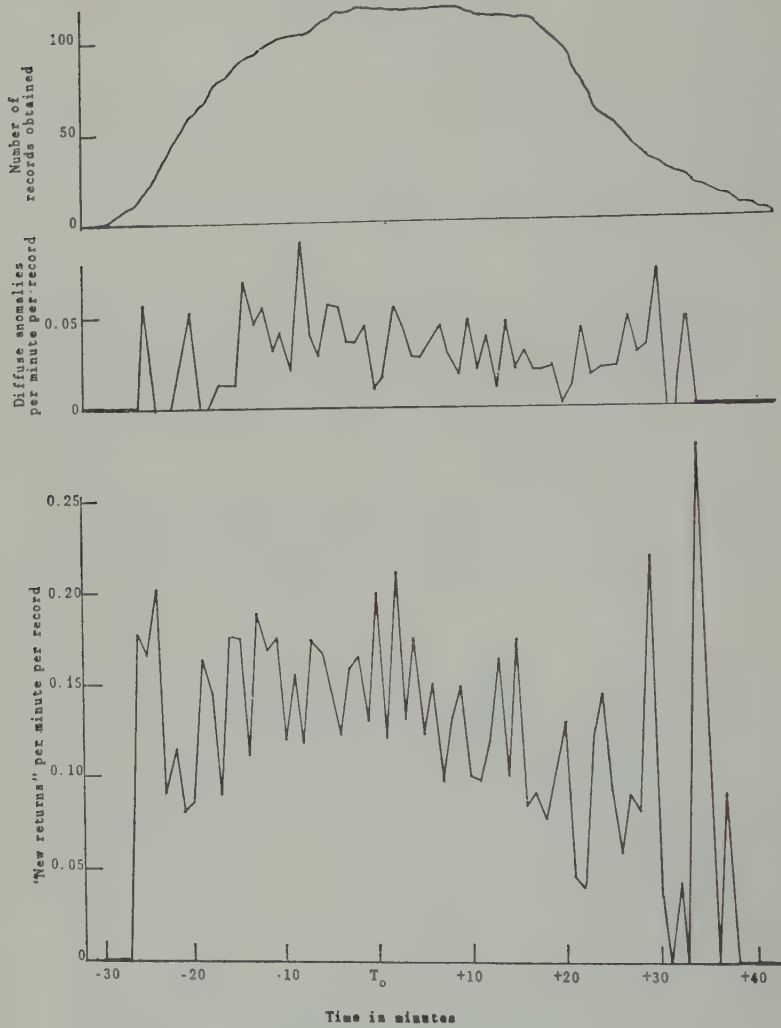


Fig. 7. Amplitude anomalies vs. T_0 time.

satellite's moving coordinate system. If the satellite were causing detectable effects which persisted after its passage, the anomalies thus plotted would be expected to show a trend to group behind the satellite. If it is assumed that the satellite, through some mechanism, caused effects which preceded it, a similar grouping would be expected ahead of the satellite. If it is assumed that the satellite caused no effect whatsoever, random grouping would be expected, with a concentration in the immediate vicinity of the satellite due to the tendency of the data analyst to downgrade or reject returns of poor correlation in range and time.

The centroid of the entire array of anomalies is shown on the plot. It seemed strange that this centroid should be so far to the right of the satellite track. Consequently the geometry of all 139 runs was analyzed to find where the centroid of the anomalies would have been if the satellite had produced no effect. As shown on the plot the proximity of this to the actual centroid indicates that a disproportionate number of experimental setups tended to look to the right of the satellite track.

The density of anomalies in the vicinity of the satellite in Figure 4 was compared with random density found during the last phase of data

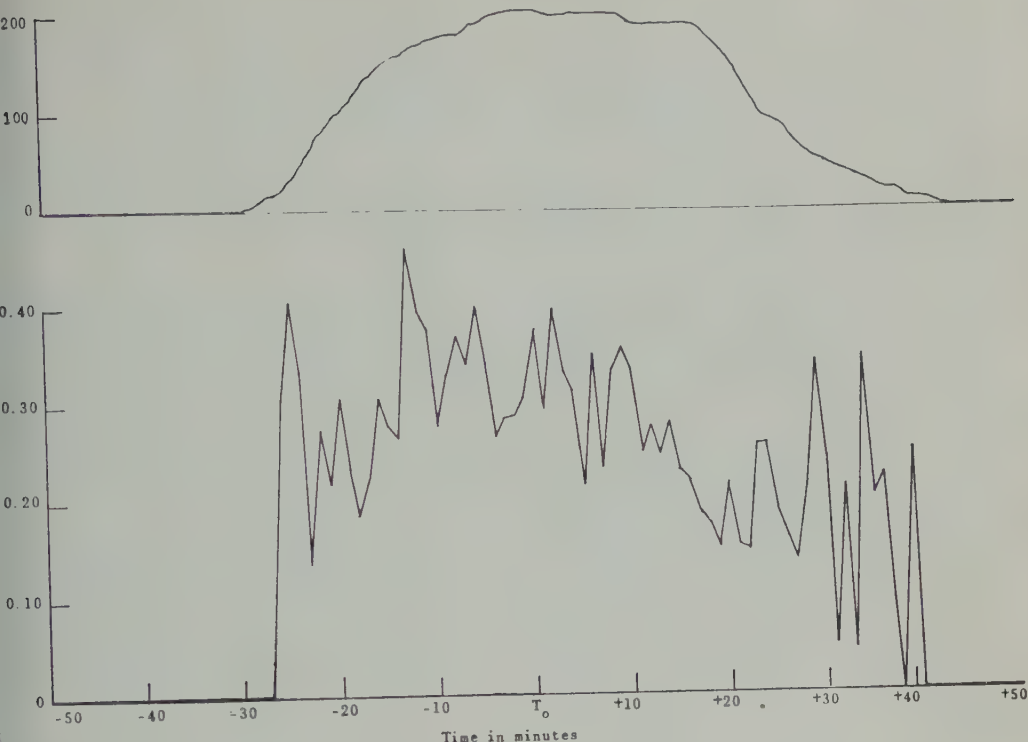


Fig. 8. Total anomalies vs. T_0 time.

duction and was bound to be in very close agreement to it.

In Figure 5 are plotted the positions of the satellite's nearest approach to the radar during every record. Passages that occurred within a horizontal distance of less than 400 km were initially studied by utilizing the log periodic antenna aimed in the direction of the satellite track. More distant passages were studied with antennas headed within 50° of the direction of the satellite's point of nearest approach. Analysis of this pattern has not disclosed any statistically significant grouping.

Only 16 solar flares were in progress during the records obtained in the intervals before or immediately after the satellite's nearest approach. The relationship between these flares and success in locating satellite-correlated anomalies was studied, but no trend was found.

The effect of the time of day on the degree of success in locating correlated anomalies was examined. No strong trends were evident, but there was slightly greater success at approximately 2000 GMT (local noon). Examination of

all records by the second reader revealed that this was the time of day when the greatest number of natural anomalies occurred.

The success in finding anomalies was next compared with geomagnetic field instability as indicated by the K_p index from the CRPL F reports of the National Bureau of Standards. There was no apparent correlation between success and the index.

In studying the effect of the frequency used in the individual experiments, it was determined that a slightly greater number of correlated anomalies were found when using frequencies in the vicinity of 12.5 Mc/s. Analysis showed that this was due in part to an experimental technique and in part to a higher average number of anomalies encountered on records obtained near that frequency.

If it is assumed that satellites induce anomalies through some mechanism involving the geomagnetic field, it might be expected that the degree of success in the detection of such effects would be a function of antenna heading. No such trend was detected in the data.

Results of the second reading of the films by a person without access to orbital data are shown in Figures 6, 7, and 8. The only form of artificial correlation the reader could have inserted would have been a tendency to find fewer anomalies near the beginning and ending times of individual records. This is due to the difficulty in identifying anomalies in the absence of records of the time immediately preceding or following the portion of the film being examined. As a consequence, there might be a slight tendency for the average number of anomalies to follow the general shape of the curve showing the number of records in effect. This tendency is present on some of the graphs.

Except for the effect discussed above, the data do not reveal any significant discontinuities. Since the satellites traveled at various velocities near 400 km per minute, one would hope to find peaks in the graphs within 1 or 2 minutes of T_0 . An attempt to locate T_0 on the curves without reference to the time scale or knowledge of its approximate location will lead to selection of various points within about a 25-minute interval on the graphs.

CONCLUSIONS

During this experiment a large number of backscatter anomalies were detected which correlated in varying degrees with passage of the satellites studied. Analysis of the entire group of records leads to the conclusion that the correlated anomalies would have been seen even in the absence of satellites. No aspect of the detected effects exhibited a consistent character from run to run. It is possible that a few anomalies seen might have been caused by satellite passage, but the number must be small enough to have a negligible effect on the type of statistical studies described. Such effects, if they exist, must be considered undetectable, since they could not be identified as unnatural in origin.

Thus, we may conclude that passage of the

satellites studied in this experiment causes redistribution in upper atmospheric ionization large enough to be detected by backscatter radar techniques such as were used here.

We may also conclude that no ionization discontinuity traveling with (or formed in the vicinity of) these satellites was of sufficient magnitude to be directly detected by this radar equipment.

Recent theories about satellite-induced effects have suggested that ionized clouds may form and then detach and move off on tracks independent of the satellite orbit [e.g., *Singer, 1961*]. Should such clouds have been present at distances of the order of 500 km from the satellite, it is unlikely that they would have been identified by the technique employed here.

Acknowledgments. We wish to thank the National Aeronautics and Space Administration and the Space Track Control Center for orbital predictions, and the Smithsonian Institution for ephemeris data on 1958 δ_2 .

This work supported by the Office of Naval Research under contract NONR 225 (33).

REFERENCES

- Cottony, H. V., and J. R. Johler, Cosmic radio noise in the VHF band, *Proc. IRE*, 40(9), 1958, 1952.
- International Telephone and Telegraph Corp., *Reference Data for Radio Engineers*, 4th ed., 1956.
- Kraus, J. D., and R. C. Higgy, The relation of the satellite ionization phenomenon to the radiation belts, *Proc. IRE*, 48(12), 2027, 1960.
- Pawsey, J. L., and R. N. Bracewell, *Radio Astronomy*, Oxford at the Clarendon Press, 1955.
- Singer, S. F., Paper presented at the March, 1961 American Astronautical Society Symposium on the Interactions of Space Vehicles with an Ionized Atmosphere, The Macmillan Co., New York, 1961.
- Van der Ziel, A., *Noise*, Prentice-Hall, Englewood Cliffs, New Jersey, 1954.

(Manuscript received April 20, 1961;
revised July 19, 1961.)

Expected Influence of a Localized Change of Ionosphere Height on VLF Propagation

JAMES R. WAIT

*Central Radio Propagation Laboratory, National Bureau of Standards
Boulder, Colorado*

Abstract. An approximate analysis is carried out for mode propagation in the earth-ionosphere waveguide with the height of the boundary varying with distance. Particular attention is paid to the phase anomaly produced by a localized depression of ionosphere height.

Introduction. The ionization produced by active clouds located in the stratosphere has been examined by *Pierce* [1961]. He has related the contours for the rate of electron production and, from these, has estimated contours of equal electron density. Such a situation arises when a nuclear explosion occurs in the stratosphere. At the instant of ionization, large quantities of energetic particles are released. These are capable of producing considerable ionization at heights below 100 km.

Pierce [1961] has suggested that to study this problem at VLF we should imagine the earth-ionosphere waveguide to have a height varying with distance. A typical depression of height might extend over as much as 1000 km.

With the above motivation we shall outline a very simple approach to the VLF propagation problem, since this aspect was not treated by *Pierce*. For present purposes we shall assume that the ionosphere is a sharply bounded ionized medium.

Choosing a rectangular coordinate system, the medium is represented by the plane $z = 0$ and the height of the lower edge of the ionosphere is given by $z = h(x)$, being a function of x only. We will now assume that $dh(x)/dx$ is always small compared with unity. Consequently, a wave propagating in such a guide does not change its form. Also, conversion of energy from one mode to another is neglected. Some justification for these assumptions can be found in the literature on multimode propagation in slightly irregular waveguides in the microwave region (*Wargan*, 1950; *Solymar*, 1959). The author [*Wait*, 1960] has also given some attention to the propagation in a linear transition region connecting two parallel-plate waveguides of differing

widths. The mode conversion in this system was calculated using an approximate technique. It was shown that, when the angle subtended by the plane bounding surfaces of the transition was small, the mode conversion was negligible and furthermore the field pattern across the guide passed smoothly from one waveguide region to the other.

The approximate solution. In view of the comments in the preceding paragraphs it is suggested that the equations already developed for VLF mode propagation [*Wait*, 1960] can be generalized to a variable ionosphere height. Thus, for propagation in the positive x direction, the vertical electric field is proportional to

$$\left[\frac{1}{h(x)} \right]^{1/2} \sum_n A_n \exp \left[-ik \int^x S_n(x) dx \right] \cos [kC_n(x)z] \quad (1)$$

where A_n are coefficients independent of x and z , $k = 2\pi/\text{wavelength}$, and $S_n(x)$ and $C_n(x)$ are dimensional factors analogous to the S_n and C_n occurring in the conventional mode theory for a constant height of the ionosphere. In the present case, the modal equation may be written

$$\left(\frac{C(x) - \Delta_i}{C(x) + \Delta_i} \right) \left(\frac{C(x) - \Delta_g}{C(x) + \Delta_g} \right) e^{-i2kh(x)C(x)} = e^{-i2\pi n} \quad (2)$$

where n is an integer, where

$$\Delta_g = (i^{1/2} e^{i\pi/4})$$

and

$$\Delta_i = \frac{[C^2(x) - (i/L)]^{1/2}}{1 - (i/L)}$$

where

$$G = \epsilon_0 \omega / \sigma_g \quad \sigma_g = \text{ground conductivity}$$

and

$$L = \frac{\omega}{\omega_r} \quad \omega_r = \frac{(\text{plasma frequency})^2}{\text{collision frequency}}$$

For purposes of this analysis, G and L are assumed to be independent of x .

The particular (complex) values of $C(x)$ that satisfy equation 2 are denoted $C_n(x)$. Then $S_n(x)$ is obtained from

$$S_n(x) = [1 - C_n^2(x)]^{1/2}$$

An approximate solution of equation 2 can be found under the assumption that

$$L \quad \text{and} \quad 1/L \gg (\hat{C}_n)^2$$

and

$$G \ll (\hat{C}_n)^2$$

where

$$\hat{C}_n = \frac{\pi(n - \frac{1}{2})}{kh(x)}$$

Thus, following an earlier analysis [Wait, 1957, 1960],

$$\begin{aligned} \text{Re } S_n(x) &\cong \hat{S}_n + \frac{1}{\sqrt{2k} h(x) \hat{S}_n} \\ &\cdot \left[\hat{C}_n^2 \left(\sqrt{L} - \frac{1}{\sqrt{L}} \right) + \sqrt{G} \right] \end{aligned} \quad (3)$$

and

$$\begin{aligned} \text{Im } S_n(x) &\cong -\frac{1}{\sqrt{2k} h(x) \hat{S}_n} \\ &\cdot \left[\hat{C}_n^2 \left(\sqrt{L} + \frac{1}{\sqrt{L}} \right) + \sqrt{G} \right] \end{aligned} \quad (4)$$

where

$$\hat{S}_n = (1 - \hat{C}_n^2)^{1/2}$$

The attenuation of a mode for propagation from x_1 to x_2 is then given by

$$- \text{Im} \int_{x_1}^{x_2} k S_n(x) dx \quad \text{nepers}$$

Once the functional form of $h(x)$ is known, the integration can be carried out to give an explicit

result. The phase can be treated in a similar manner.

Application to a special form of height depression. A specific example is now carried out to illustrate the nature of the phenomenon. We set

$$\begin{aligned} \frac{kh(x)}{(n - \frac{1}{2})\pi} &= a - b \cos \frac{\pi}{2} \frac{x}{x_0} \\ &\text{for } -x_0 < x < x_0 \\ &= a \quad \text{for } |x| > x_0 \end{aligned}$$

where

$$a = \frac{kh_0}{(n - \frac{1}{2})\pi} = \frac{1}{\hat{C}_n} \quad \text{and} \quad b = \frac{k \Delta h}{(n - \frac{1}{2})\pi}$$

In the above, h_0 is the constant height of undisturbed ionosphere and Δh is the maximum depression of the disturbed region. The situation is illustrated in Figure 1.

The attenuation P_n , in nepers, for the part of the path from $-x_0$ to x is given by

$$\begin{aligned} P_n &= -k \text{Im} \int_{-x_0}^x S_n(x) dx \\ &= k \delta_n \int_{-x_0}^x \frac{1}{[a - b \cos (\pi/2)(x/x_0)]^3} dx \end{aligned}$$

where

$$\delta_n = \frac{\sqrt{L} + (1/\sqrt{L})}{\sqrt{2}(n - \frac{1}{2})\pi}$$

The integration can be carried out in closed form but a simplification can be made if b is assumed to be small with respect to a , or $\Delta h \ll h_0$. Then

$$\begin{aligned} P_n &\cong \frac{k \delta_n}{a^3} \int_{-x_0}^x \left[1 + 3 \frac{b}{a} \cos \left(\frac{\pi}{2} \frac{x}{x_0} \right) \right. \\ &\quad \left. + 6 \left(\frac{b}{a} \right)^2 \cos^2 \left(\frac{\pi}{2} \frac{x}{x_0} \right) + \dots \right] dx \\ &\cong k \frac{\delta_n}{a^3} (x + x_0) \left\{ 1 + \frac{6}{\pi} \frac{\Delta h}{h_0} \frac{x_0}{x + x_0} \right. \\ &\quad \cdot \left(1 + \sin \frac{\pi}{2} \frac{x}{x_0} \right) \\ &\quad \left. + 6 \left(\frac{\Delta h}{h_0} \right)^2 \left[\frac{1}{2} + \frac{x_0}{2(x + x_0)\pi} \right. \right. \\ &\quad \left. \left. \cdot \sin \left(\frac{\pi}{2} \frac{x}{x_0} \right) \right] + \dots \right\} \end{aligned}$$

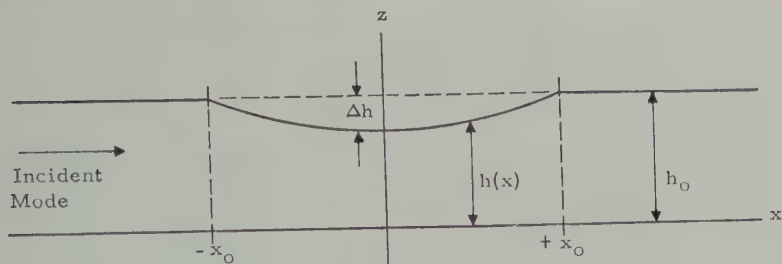


Fig. 1. The waveguide model showing the idealized perturbation of the upper boundary.

second and higher terms, inside the square brackets of the preceding expression, represent fractional increase of the attenuation. For integration across the whole disturbed region, $-x_0$ to x_0 , we see that

$P_n^{(0)} \left[1 + \frac{6}{\pi} \frac{\Delta h}{h_0} + 3 \left(\frac{\Delta h}{h_0} \right)^2 + \dots \right] \quad (8)$

$P_n^{(0)}$ is the attenuation for the same path in the absence of a disturbance. In the above

$= 2kx_0 \frac{\delta_n}{a^{\frac{2}{3}}}$

$\frac{\sqrt{2}x_0}{h_0} \left(\left(n - \frac{1}{2} \right) \pi \right)^2 \left(\sqrt{L} + \frac{1}{\sqrt{L}} \right) \quad (9)$

is rather interesting to note that the square-bracket term in equation 8 depends only on the fractional depression of the ionosphere reflecting at.

Of greater interest is the influence of the disturbed region on the phase. The electrical path length, for mode n , from $-x_0$ to x is previously given by

$\Phi_n = k \int_{-x_0}^x \text{Re } S_n(x) \, dx \quad (10)$

Using equation 3, and remembering that $1/a \delta_n$ is small compared with unity, we find that

$\approx k(x + x_0) - \frac{k}{2} \int_{-x_0}^x \frac{dx}{(a - b \cos y)^2} + k \delta_n' \int_{-x_0}^x \frac{dx}{(a - b \cos y)^3} \quad (11)$

$= \frac{\pi}{2} \frac{x}{x_0} \quad \text{and} \quad \delta_n' = \frac{\sqrt{L} - (1/\sqrt{L})}{\sqrt{2} \left(n - \frac{1}{2} \right) \pi}$

Immediately it can be noted that the latter

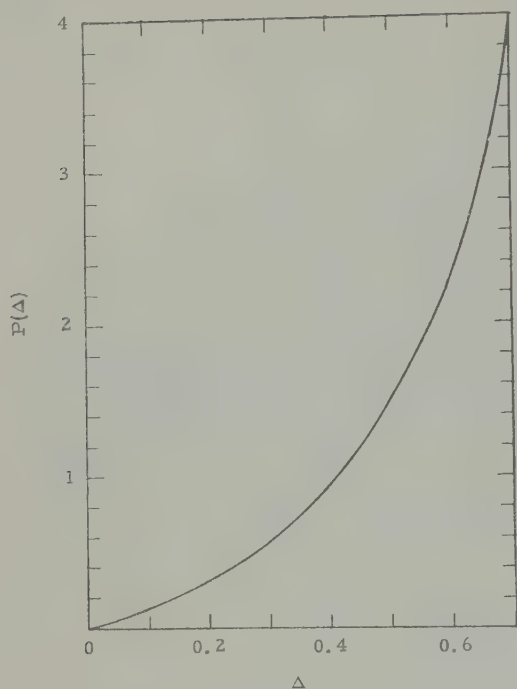
integral is of the same form as discussed above. The preceding integral is of a similar form, but in this case we shall not assume that $b \ll a$ and it may be evaluated in closed form by noting that

$$\int \frac{dy}{(a - b \cos y)^2} = \frac{1}{a^2 - b^2} \left[\frac{b \sin y}{(a - b \cos y)} + \frac{2a}{\sqrt{a^2 - b^2}} \tan^{-1} \frac{\sqrt{a^2 - b^2} \tan(y/2)}{a - b} \right] \quad (12)$$

Consequently

$$\Phi_n = k(x + x_0) - \frac{kx_0}{\pi(a^2 - b^2)} \left\{ \frac{b \sin \frac{\pi}{2} \frac{x}{x_0}}{a - b \cos \frac{\pi}{2} \frac{x}{x_0}} + \frac{2a}{\sqrt{a^2 - b^2}} \tan^{-1} \frac{\sqrt{a^2 - b^2} \tan \left(\frac{x}{x_0} \frac{\pi}{4} \right)}{a - b} + \frac{b}{a} + \frac{2a}{\sqrt{a^2 - b^2}} \tan^{-1} \sqrt{\frac{a + b}{a - b}} \right\} + k \frac{\delta_n'}{a^{\frac{2}{3}}} (x + x_0) \left\{ 1 + \frac{6}{\pi} \frac{b}{a} \frac{x_0}{x + x_0} \cdot \left(1 + \sin \frac{\pi}{2} \frac{x}{x_0} \right) + 6 \left(\frac{b}{a} \right)^2 \cdot \left[\frac{1}{2} + \frac{x_0}{2(x + x_0)\pi} \sin \pi \frac{x}{x_0} \right] + \dots \right\} \quad (13)$$

Of special interest is the total 'phase anomaly' for propagation of a mode across the disturbed

Fig. 2. The function $P(\Delta)$.

region (from $-x_0$ to x_0). This quantity, denoted PA , is defined by

$$PA = \Phi_n - \Phi_n]_{\Delta h=0} \quad (14)$$

Using equation 13, it readily follows that

$$PA = -\frac{kx_0}{a^2} [P(\Delta) - \epsilon] \quad (15)$$

where

$$P(\Delta) = \frac{2}{\pi(1-\Delta^2)} \left[\Delta + \frac{2}{\sqrt{1-\Delta^2}} \cdot \tan^{-1} \sqrt{\frac{1+\Delta}{1-\Delta}} \right] - 1 \quad (16)$$

and

$$\epsilon = \frac{2\delta_n'}{a} \left[\frac{6}{\pi} \Delta + 3\Delta^2 + \dots \right] \quad (17)$$

where

$$\Delta = b/a = \Delta h/h_0$$

In most cases of interest ϵ can be neglected in comparison with $P(\Delta)$. Thus the phase anomaly is proportional to $P(\Delta)$. This dimensionless

factor is shown plotted in Figure 2 as a function of Δ from 0 to 0.7. When $\Delta \ll 1$ it is seen that

$$P(\Delta) \cong 4\Delta/\pi$$

It may be noted that $P(\Delta)$ does not depend on the mode number n or the distance $2x_0$.

For the dominant mode, $n = 1$, it is seen that

$$\begin{aligned} PA &\cong -\frac{\pi}{8} \frac{(x_0/\lambda)^2}{(h_0/\lambda)^2} P(\Delta) \quad \text{radians} \\ &\cong -22.5 \frac{(x_0/\lambda)^2}{(h_0/\lambda)^2} P(\Delta) \quad \text{degrees} \end{aligned} \quad (18)$$

The negative sign preceding these expressions signifies that the PA is a phase advance.

To illustrate the order of magnitude of this quantity, a typical set of values is chosen following the suggestions of Pierce. These are

$$\begin{aligned} 2x_0 &= 800 \text{ km} & h_0 &= 80 \text{ km} & \Delta h &= 20 \text{ km} \\ f &= 15 \text{ kc/s} & (\text{i.e., } \lambda &= 20 \text{ km}) \end{aligned}$$

A simple calculation gives

$$PA \cong -11.8 \quad \text{degrees}$$

Influence of earth curvature. In the preceding discussion the influence of earth curvature has not been considered. To account for this fully leads to great complications [Wait and Spies, 1960]. However, a first-order correction for the phase can be obtained rather simply if it is assumed that the bending of the waveguide does not change the phase velocity relative to the center of the guide. Thus, for a finite value of the radius a_e of the earth,

$$\text{Re } S_n \cong \text{Re } S_n]_{a_e \rightarrow \infty} \times \left(\frac{2a_e + h(x)}{2a_e} \right) \quad (19)$$

Since $h/a_e \ll 1$, and remembering that $\hat{C}_n \ll 1$, it is permissible to replace the above by

$$\text{Re } S_n \cong \text{Re } S_n]_{a_e \rightarrow \infty} + \frac{h(x)}{2a_e} \quad (20)$$

Thus, equation 10 is replaced by

$$\begin{aligned} \Phi_n &= k \int_{-x_0}^x \text{Re } S_n(x) dx \\ &\quad + \frac{k}{2a_e} \int_{-x_0}^x h(x) dx \end{aligned} \quad (21)$$

When $h(x) = h_0 - \Delta h \cos(\pi x/2x_0)$ as in the previous example, it readily follows that, for

$$= \Phi_n]_{x_0 \rightarrow \infty} + \frac{k}{2a_s} (x + x_0) h_0 - \frac{kx_0}{\pi} \frac{\Delta h}{a_s} \left[1 + \sin \left(\frac{\pi x}{2x_0} \right) \right] \quad (22)$$

which is valid for the interval $-x_0 < x < x_0$. Using this approach, the total phase anomaly given by

$$= -\frac{kx_0}{a_s^2} [P(\Delta) - \epsilon] - \frac{2kx_0}{\pi} \frac{\Delta h}{a_s} \quad (23)$$

here, as before,

$$a = kh_0 / [(n - \frac{1}{2})\pi] = 1/\hat{C}_n$$

The latter term, on the right-hand side of equation 23, which involved the earth curvature, can be neglected if

$$\hat{C}_n^2 \gg h_0/a_s$$

For frequencies greater than about 10 kc/s and the dominant mode (i.e., $n = 1$), it turns out that the curvature correction term is quite significant. For, $n = 1$, and neglecting ϵ ,

$$A = -\frac{\pi}{8} \frac{(x_0/\lambda)^2}{(h_0/\lambda)^2} P(\Delta) - 4 \left(\frac{x_0}{\lambda} \right) \frac{h_0}{a_s} \Delta \quad (24)$$

where $\Delta = \Delta h/h_0$. Taking the same example as before (i.e., $f = 15$ kc/s, $2x_0 = 800$ km, $h_0 = 80$ km, $\Delta h = 20$ km), this leads to

$$PA = -11.8^\circ - 14.4^\circ = -26.2^\circ$$

This is an order of magnitude relatively easy to measure.

Conclusion. Although the present analysis is based on a highly idealized model with a number of simplifying assumptions, it does illustrate the great sensitivity of VLF phase to perturbations in the lower ionosphere. Further theoretical investigations of this problem, using more elaborate approaches, are continuing.

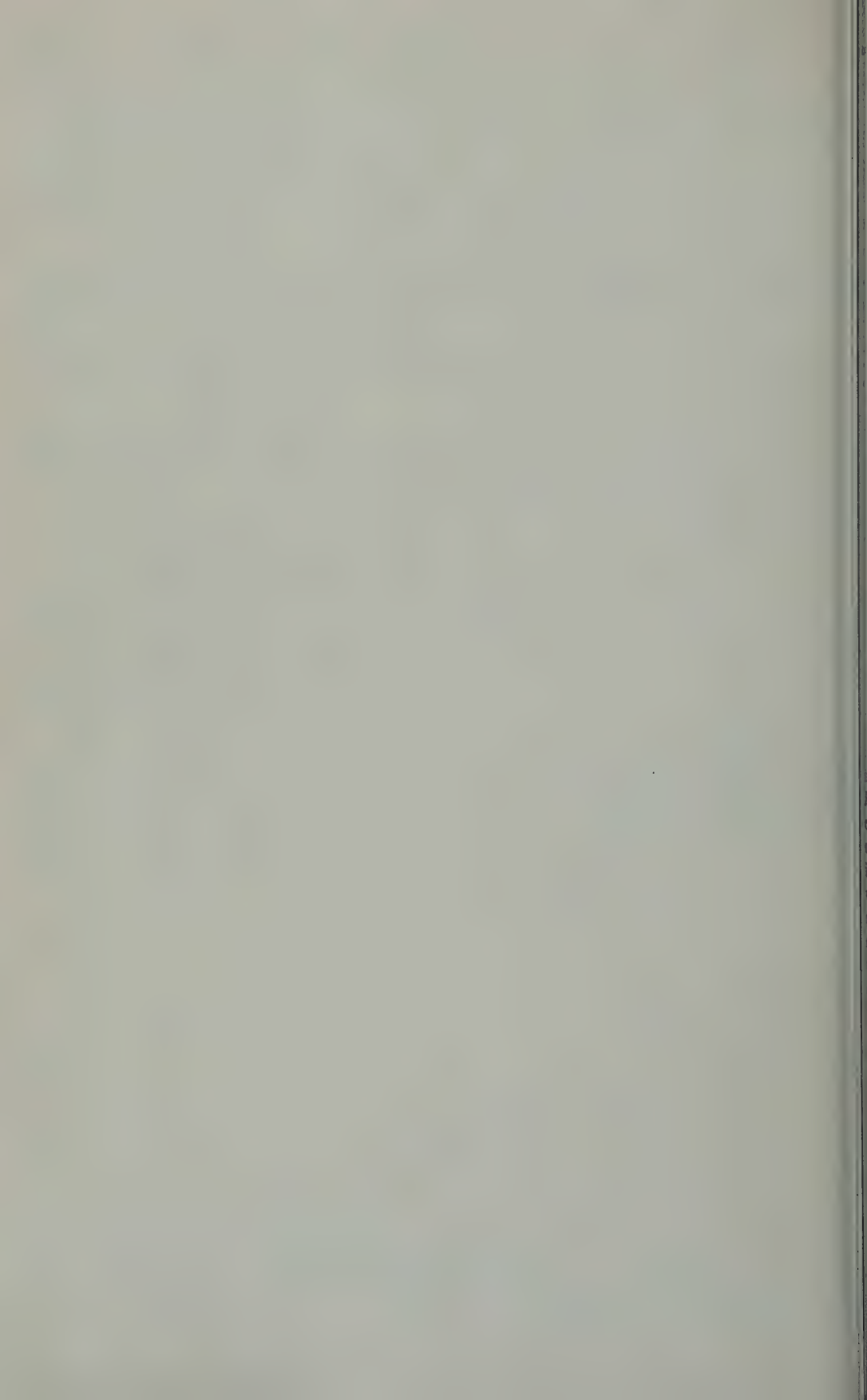
Acknowledgment. I should like to thank Mrs. Alyce Conda for checking the algebra and Mrs. Eileen Brackett for assistance in preparing the paper.

The research reported in this paper has been sponsored by the Advanced Research Projects Agency, Washington 25, D. C.

REFERENCES

- Morgan, S. P., Mode conversion losses in transmission of circular electric waves through slightly non-cylindrical guides, *J. Appl. Phys.*, **21**, 329-338, 1950.
- Pierce, E. T., Ionization below 100 km due to radioactive clouds, *Stanford Research Inst. Tech. Rept.*, Project 3097, March 1961.
- Solymar, L., Overmoded waveguides, *Electronic and Radio Engineer*, **36**, 426-428, 1959.
- Wait, J. R., The mode theory of VLF ionospheric propagation for finite ground conductivity, VLF Symposium paper 6, Boulder, January 1957, *Proc. IRE*, **45**(6), 760-767, 1957.
- Wait, J. R., Terrestrial propagation of very-low-frequency radio waves, a theoretical investigation, *J. Research NBS*, **64D** (Radio Prop.), (2), 153-204, 1960.
- Wait, J. R., A preliminary analysis of VLF mode propagation for a variable ionosphere height, *NBS Rept.* 6770, June 5, 1961.
- Wait, J. R., and K. Spies, Influence of earth curvature and the terrestrial magnetic field on VLF propagation, *J. Geophys. Research*, **65**, 2325-2331, 1960.

(Manuscript received July 5, 1961.)



The Interpretation and Synthesis of Certain Spread- F Configurations Appearing on Equatorial Ionograms

WYNNE CALVERT AND ROBERT COHEN

*Central Radio Propagation Laboratory, National Bureau of Standards
Boulder, Colorado*

Abstract. On ionograms obtained near the magnetic equator the rectangular configuration called 'equatorial spread F ' arises from scattering in the vertical plane (passing through the ionosonde) normal to thin, magnetic field-aligned irregularities located at or beneath the base of the F layer [Cohen and Bowles, 1961]. It is shown in the present paper that some strikingly different spread- F configurations on equatorial ionograms are due to irregularities of the same kind, but embedded in the F layer. Since these other configurations result from similar irregularities, the category of 'equatorial spread F ' is generalized to include them. An 'ionogram' corresponding to scattering from an individual 'equatorial spread F ' irregularity can be calculated. A composite ionogram resulting from a number of such scatterers can then be synthesized by superposition. By appropriate distributions of scattering centers in the east-west plane, many features of the 'equatorial spread- F ' configurations observed on Huancayo ionograms can be simulated. This process of ionogram synthesis constitutes a new technique for the determination of (1) the height of patches of small-scale irregularities in the F region; (2) the horizontal distance of the patches from the ionosonde in the vertical east-west plane; (3) the thickness of the patches; (4) their east-west extent; and (5) the east-west component of their drift velocity. The patches are measured to have an east-west extent of up to 300 km and a thickness between 10 and 50 km. Their drift velocity is eastward, gradually decreasing from 200 to 100 meters/sec after 2100 hours local time.

INTRODUCTION

Purpose. In this paper, basic information about ionospheric irregularities in the equatorial region will be utilized for the interpretation of certain spread- F configurations observed on equatorial ionograms. A propagation model involving refraction in the intervening ionosphere will be used to trace rays between the ionosonde and the irregularities, in order to explain and synthesize the observed echo configurations. In the process of simulating such ionograms, considerable information can be obtained about the aggregations of irregularities giving rise to equatorial spread F .

The characteristics of spread echoes on ionograms depend upon (a) the nature of the scattering irregularities, and (b) the propagation of rays from the ionosonde to and from the scatterers. The study of spread echoes near the magnetic equator achieves a simplification of their interpretation, because the spread- F irregularities are known to be elongated along the Earth's magnetic field. Thus, at the magnetic equator the propagation is confined to the vertical magnetic east-west plane passing through

the ionosonde (since for an echo to be obtained from an elongated irregularity a wavefront must arrive at the irregularity parallel to the elongation).

Background. A series of observations has been performed [Cohen and Bowles, 1961] to augment the information available from the ionosonde at Huancayo, Peru ($12^{\circ}3'S$, $75^{\circ}20'W$: magnetic dip, $2^{\circ}N$). Two classes of spread- F irregularities have been recognized in these studies, on the basis of their ability or inability to scatter 50-Mc/s radio waves obliquely. The 'scattering' property was associated with a rectangular spread- F configuration, a spread in range relatively independent of frequency, described as 'equatorial spread F ' (Fig. 1). The 'nonscattering' irregularities produced a configuration having a spread in frequency relatively independent of range, described as 'temperate-latitude spread F ' (Fig. 2), due to its similarity in appearance to that observed in temperate latitudes. The terms 'range spreading' and 'frequency spreading,' respectively, have been used to describe such spread- F configurations [McNicol, Webster, and Bowman, 1956].

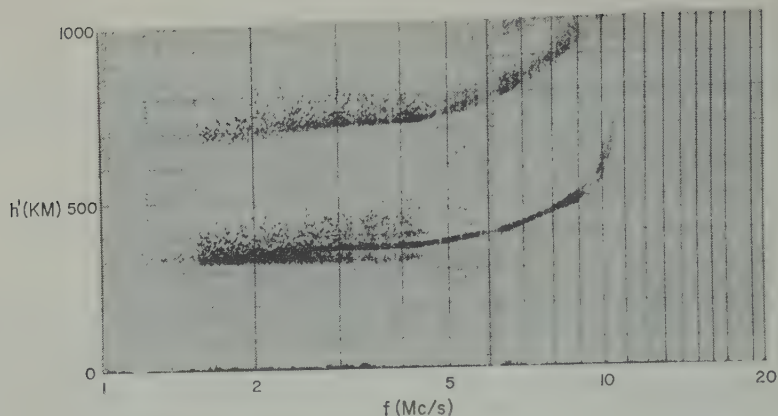


Fig. 1. A rectangular 'equatorial spread-F' configuration observed at Huancayo, Peru, 2100 EST, October 29, 1960.

Certain other configurations (Figs. 3 and 4) were apparently anomalous in that they were associated with scatter propagation, yet often (e.g., Fig. 3) more closely resembled the 'temperate-latitude spread F' ' than the 'equatorial spread F' '. Analysis of these configurations motivates broadening the 'equatorial spread F' ' category to include them. (The 'temperate-latitude spread F' ' phenomenon is treated in another paper [Pitteway and Cohen, 1961].)

Characteristics of the scattering irregularities. It has been established [Cohen and Bowles, 1961] that thin, elongated irregularities occur in the nighttime equatorial F' region at heights as great as 450 km. These irregularities were demonstrated to be aligned with the earth's magnetic field and to extend 1 km or more in the field direction. In at least one direction transverse to the field

the irregularities were shown to be of 10-meter dimensions or smaller. They occur in patches having a thickness of the order of 50 km, and an east-west extent as great as 1000 km. These patches were found to be located at or as much as 100 km below the base of the F' layer.

THE MODEL

Basic considerations. The spread- F' configurations of Figures 1, 3, and 4 can be ascribed to single scattering in the vertical east-west plane from thin, field-aligned irregularities in the F' region. To show how these configurations arise, a model is now presented which embodies refraction in the intervening ionosphere and scattering from the irregularities (subject to an orthogonality condition).

This model is included in the general scattering

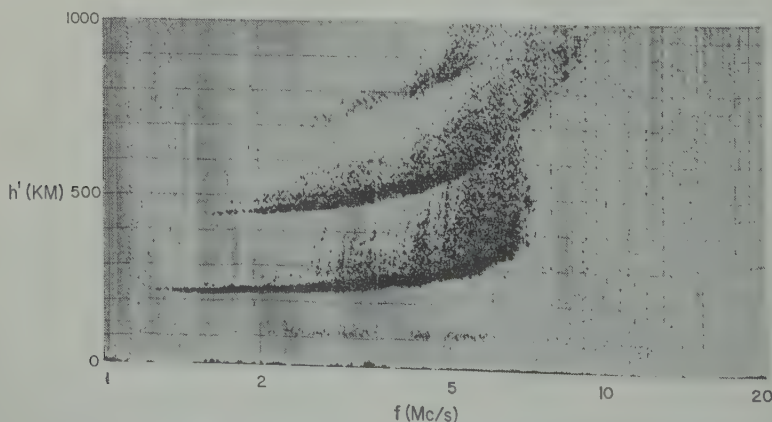


Fig. 2. The 'temperate-latitude spread F' ' configuration observed at Huancayo, 0216 EST, April 22, 1960. Note the striations visible in the spread F' .

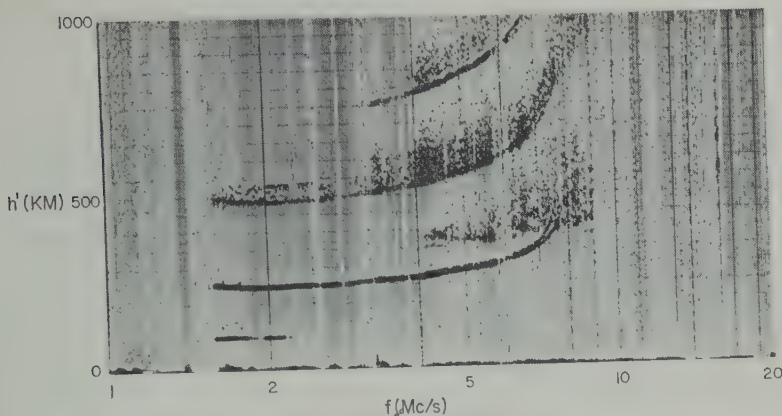


Fig. 3. An additional spread- F configuration frequently observed at Huancayo. Recorded at 2315 EST, July 5, 1960.

ries proposed by *Eckersley* [1937, 1940] and *Peterminger* [1951]. Similar models have also been proposed by *Peterson* [1951], *Gassmann* [1957], *Renau* [1960], and *Booker* [1961].

The present model considers the relatively simple geometry operating in the equatorial case, as indicated by previous experimentation concerning the nature of the scattering irregularities. The conditions involved here in the simulation of spread ionograms by ray tracing are that (a) the ray-tracing approach is valid (i.e., there are only slow variations of electron density); (b) propagation is confined to the vertical east-west plane; (c) only the ordinary ray is considered; (d) the ionosphere is horizontally stratified; (e) the irregularities scatter isotropically in the east-west plane; and (f) a single-scattering approximation is valid.

As has been established by *Cohen and Bowles* [1961], the irregularities associated with spread F at the magnetic equator occur in the F region and are elongated along the earth's magnetic field; that is, they are horizontal and oriented north-south. Thus, thin irregularities in the equatorial F region observed from a point on the magnetic equator will produce scatter echoes only in the vertical east-west plane of orthogonality. This result is the basis for condition (b). Figure 5 shows the ray geometry in this east-west plane, for an ionosonde at A and an elongated irregularity with cross section at B .

At the equator, propagation in the vertical east-west plane is transverse to the magnetic field lines; hence the dispersion relation for the ordinary wave is identical to that for the field-free case, namely,

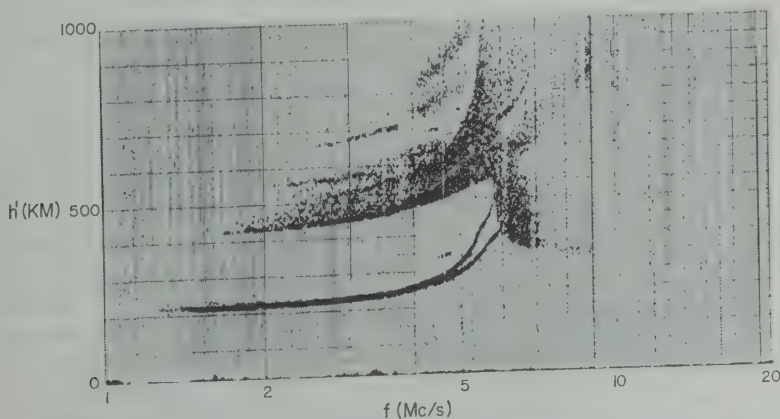


Fig. 4. Another spread- F configuration often observed at Huancayo after midnight. Recorded at 0414 EST, April 28, 1960.

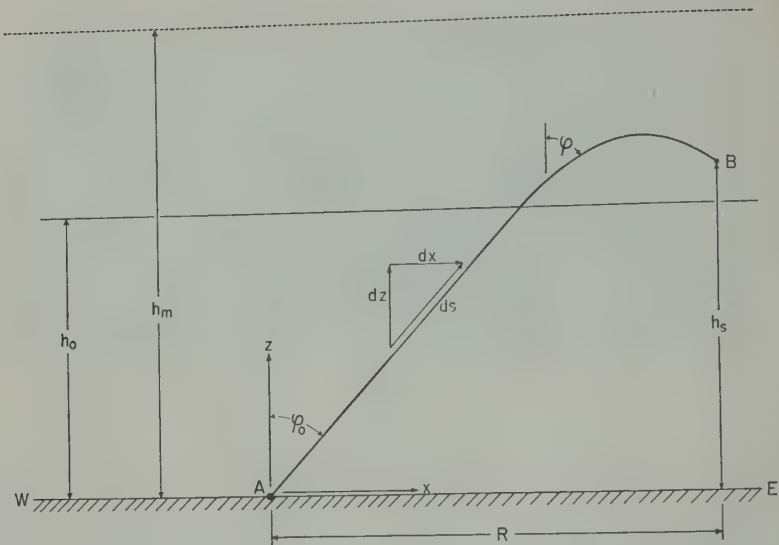


Fig. 5. Ray-trajectory to an irregularity at point B in the vertical east-west plane from an ionosonde at point A . The height of the irregularity is h_s , its ground distance from the ionosonde is R , and the take-off angle of the ray is ϕ_0 . The 'base' of the F layer is at height h_0 , defined as the height of the associated electron-density distribution at which refraction becomes important. The 'peak' of this F layer is at height h_m , and refers to the height of maximum electron density.

$$n^2 = 1 - f_N^2 / f^2 \quad (1)$$

where n is the index of refraction, f_N is the plasma frequency, and f is the wave frequency. Furthermore, this relation is also approximately true for the extraordinary wave, provided that the propagation is sufficiently oblique [Ratcliffe, 1960]. For this reason the extraordinary scatter traces are seldom resolved in 'equatorial spread F ' configurations [Pitteway and Cohen, 1961].

Conditions (a), (c), and (d) permit the use of Snell's law in calculating propagation paths. Condition (e), equivalent to regarding the irregularities as cylindrical, implies that a scatter echo may be obtained over any ray path arriving at a scatterer, or by scattering from one such ray path into another. However, the allowed ray solutions are limited to those that emanate from the ionosonde location and arrive at the irregularity. The equivalent range along such paths, h' , is then plotted vs. frequency to produce an ionogram corresponding to each irregularity. Condition (f) implies that the individual irregularities scatter independently; hence their collective scattering behavior may be treated as a superposition of the scattering from the individual scattering centers, and the respective ionograms may be superimposed.

The 'ionogram' resulting from a discrete scat-

tering center. In accordance with the above assumptions, the 'ionogram' ($h' = h'(f)$) corresponding to a given scattering center may be calculated by ray tracing. First, a scattering irregularity directly overhead will be considered. The results can then be transformed so as to be applicable for an irregularity located east or west of the ionosonde.

For the overhead case, the equivalent (vertical) range, h' , to an irregularity at height h_s is first computed as a function of frequency for the allowable ray paths. These paths and the resulting ionograms are depicted in Figures 7a and 8a for an irregularity situated, respectively, at or below the base (h_0) of the F layer, and at or above the peak of the layer (h_m). (The representative parameters of the parabolic F layer employed in these calculations were $h_0 = 200$ km; $h_m = 300$ km; the F -layer penetration frequency, $f_oF = 7.0$ Mc/s.)

For the first two height intervals (Figs. 7a and 7b), there are three possible 'round-trip' ray paths: (i) the direct path to the scatterer and return; (ii) the path reflected to the scatterer by the overlying ionosphere, and return; (iii) a clockwise or counterclockwise loop, involving the traversal of the direct path to

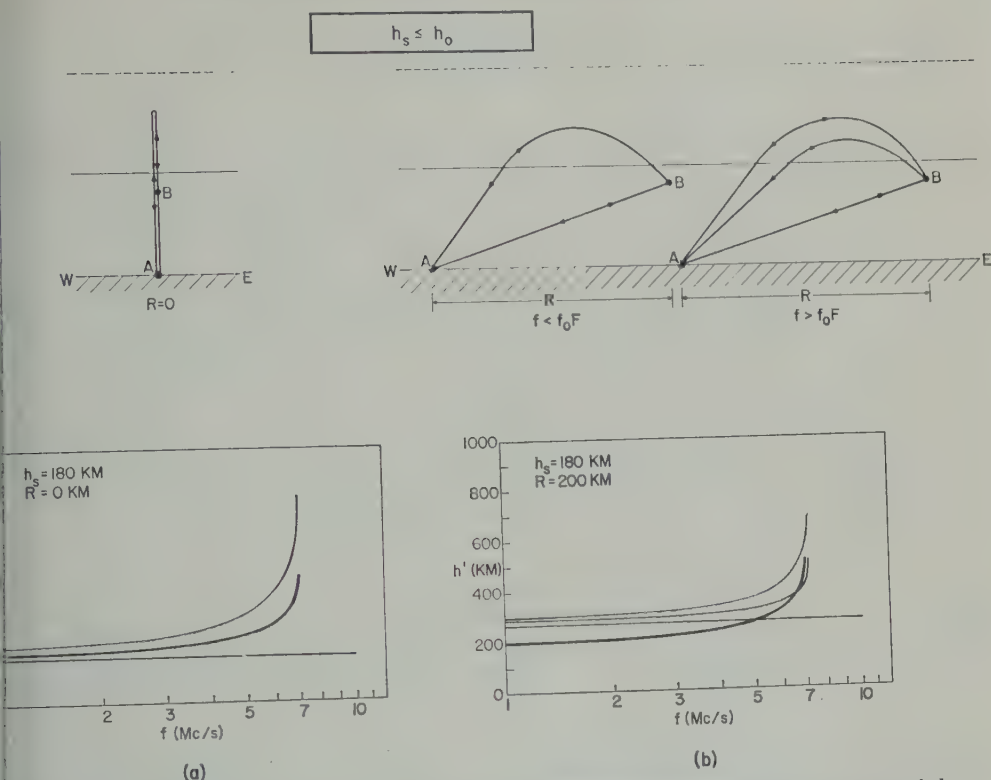


Fig. 6. Ray paths and the resulting 'ionograms' for an irregularity at or below the base of the F layer: (a) directly overhead; (b) at a ground distance of 200 km from the ionosonde. (The parabolic F layer used in this case and in the calculations of Figs. 7, 8, and 10 has its base at 200 km and its peak at 300 km, and a penetration frequency of 7.0 Mc/s.)

er and returning via the overlying ionosphere, or vice versa. The 'backscatter' modes, ii, result in the two extra traces on the ionograms of Figures 6a and 7a. These traces are below and above the main trace, respectively. The 'combination' mode, iii, produces a trace having an equivalent range given by the average of those for the mode i and mode ii traces. For an overhead scatterer, the average range coincides with the regular F trace. For the third height (Fig. 8a), i.e., for an irregularity at or above the peak of the F layer, one ray path is possible, producing a single, 'backscatter' trace at frequencies beyond the penetration frequency of the layer. This trace occurs just at the penetration frequency, where it is greatly retarded. With increasing frequency its equivalent range decreases, approaching the height, h_s , of the irregularity. As a conventional ionosonde is effective above its usual 'ceiling' at h_m when there are appro-

priately strong scattering irregularities present at greater heights.)

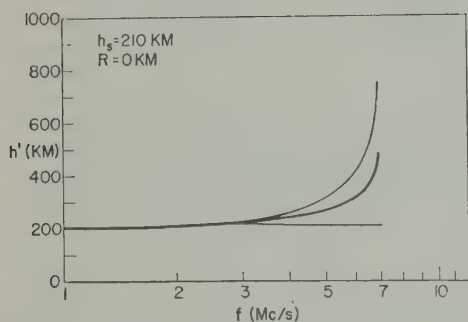
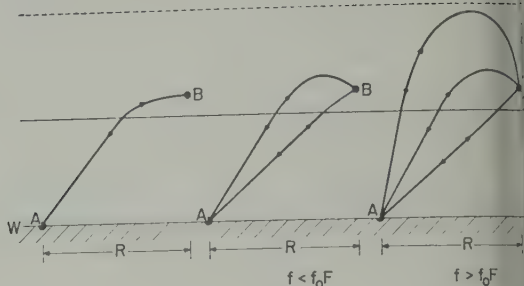
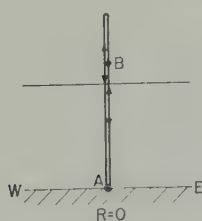
There is a point-to-point correspondence between the 'backscatter' traces for an overhead irregularity and those for an irregularity at the same height located east or west of the ionosonde (as shown in the Appendix). This correspondence is given by the transformation

$$\begin{cases} f_R = f_V(1 + R^2/h_V'^2)^{1/2} \\ h_R' = h_V'(1 + R^2/h_V'^2)^{1/2} \end{cases} \quad (2)$$

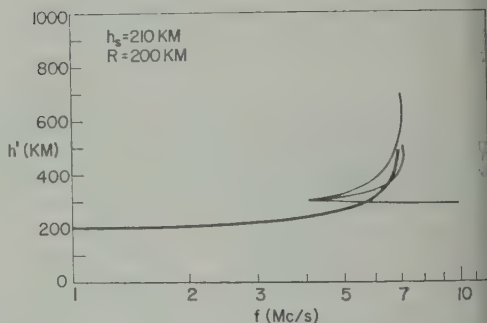
where $h_V' = h_V'(f_V)$ represents an ionogram trace produced by an overhead irregularity as a function of the vertical incidence frequency, f_V ; and $h_R' = h_R'(f_R)$ is the corresponding ionogram trace for an irregularity at the same height at a ground distance, R , east or west of the ionosonde, as a function of the frequency at oblique incidence, f_R .

This transformation is a generalization of

$$h_0 < h_s < h_m$$



(a)



(b)

Fig. 7. Ray paths and the resulting 'ionograms' for an irregularity embedded between the base and the peak of the F layer: (a) directly overhead; (b) at a ground distance of 200 km from the ionosonde.

Martyn's theorem [Martyn, 1935], where the common factor $(1 + R^2/h_v'^2)^{1/2}$ is the secant of the takeoff-angle ϕ_0 (Fig. 5). The transformation is derived in the Appendix under the conditions: (a) the oblique rays are reflected at the same height in the ionosphere as the corresponding rays at vertical incidence; (b) the scatterer is located at the same height in both the vertical and oblique cases.

Transformation 2 has the effect of projecting the 'backscatter' traces toward higher frequencies along straight lines passing through the origin (on an ionogram plotted with a *linear* rather than *logarithmic* frequency axis), as in Figure 9. Thus, for an overhead irregularity embedded between the base and the peak of the F layer (Fig. 7a) the scatter traces begin abruptly at the local plasma frequency in the vicinity of the irregularity, f_v , and at an equivalent range, h_v' , whereas for an oblique irregularity at the same

height (Fig. 7b) this 'triple point' occurs at higher frequency, f_R , and at a greater equivalent range, h_R' . Such triple points may even occur beyond the penetration frequency of the layer (Fig. 10; $h_s = 220$ km, $R = 400$ km).

'Combination' traces (mode iii) in the oblique case are still calculated by averaging the equivalent ranges, $h_R'(f)$, for the two 'backscatter' traces (modes i and ii). Furthermore, such 'combination' trace would be expected to surpass in intensity the 'backscatter' traces with which it is associated, for two reasons: there are two ray paths of equal retardation combining to produce the mode iii contribution; and the scattering at the irregularity in mode iii through an angle of less than or equal to π radians, whereas modes i and ii correspond to scattering through an angle of π radians. (The scattered intensity is presumed to be much greater, the smaller the scattering angle.)

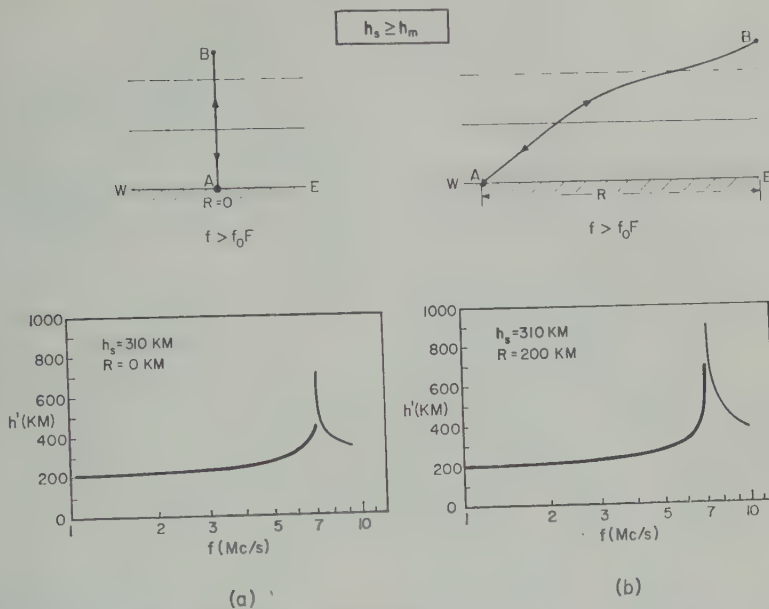


Fig. 8. Ray paths and the resulting 'ionograms' for an irregularity embedded in the F layer at above its peak electron density: (a) directly overhead; (b) at a ground distance of 200 km from the ionosonde.

for the penetration frequency of the layer, the mode ii trace for an overhead irregularity may be greatly retarded ($h'_v \rightarrow \infty$ as $f \rightarrow f_0 F$), the corresponding trace for an irregu-

larity at the same height located east or west of the ionosonde must double back, to approach an asymptote at that penetration frequency (Fig. 9). This is apparent from transformation (2) when

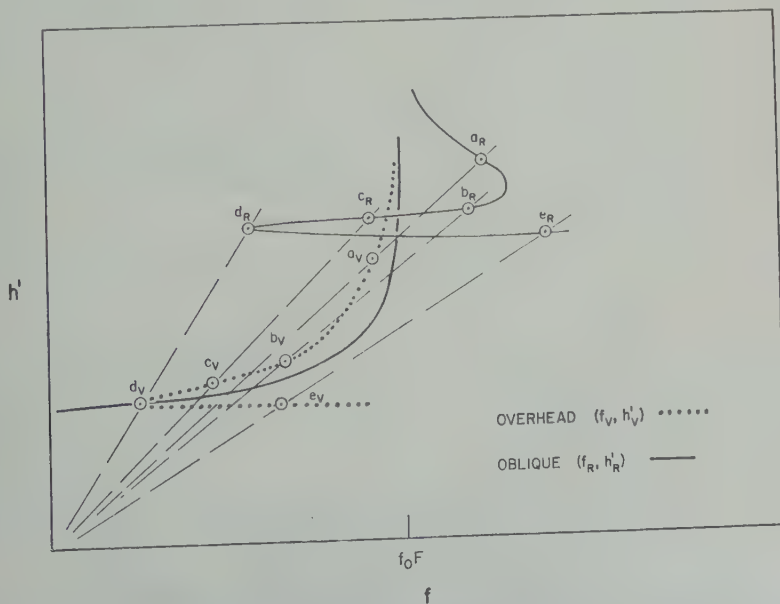


Fig. 9. Schematic transformation of the 'backscatter' traces for an irregularity embedded between h_0 and h_m (the two cases of Fig. 7).

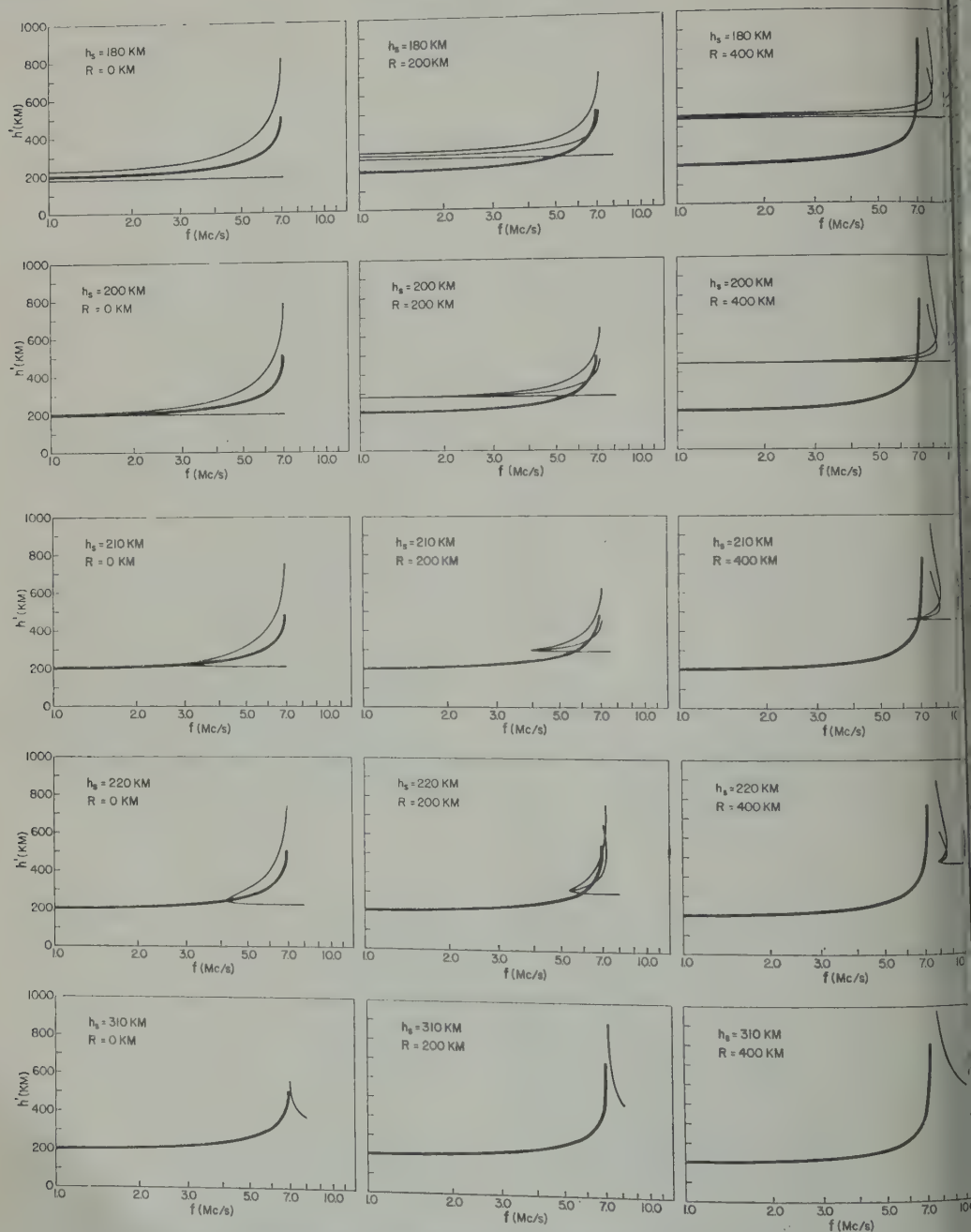
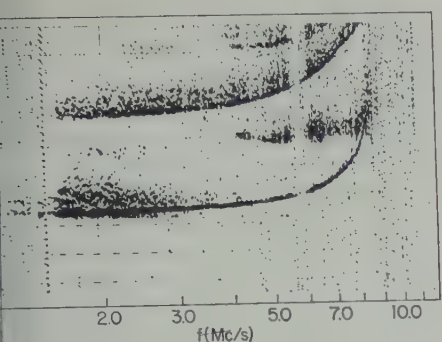


Fig. 10. Predicted 'ionograms' for irregularities at various heights, h_s , and ground distances, R , east or west of an equatorial ionosonde.

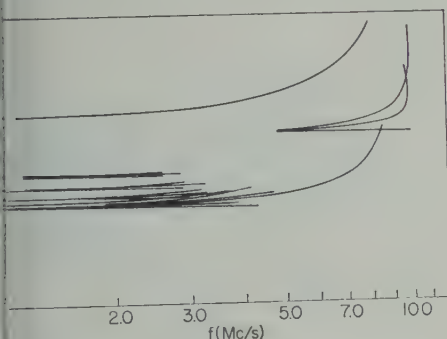
h'_v becomes large ($f_R \approx f_v \rightarrow f_0 F$; $h'_R \approx h'_v \rightarrow \infty$). This doubling-back to $f_0 F$ (Figs. 6b, 7b, 9, and 10) is a manifestation of the Pedersen mode. The 'nose' frequency in such cases thus plays

the role of a 'maximum usable frequency' (MU) between the ionosonde and the irregularity for the scatter paths involving a reflection.

Figure 10 is a collection of predicted 'ionograms' for



(a)



(b)

11. (a) An ionogram at Huancayo in which spread- F configurations occur simultaneously. (b) The simulation of this ionogram, with scattering irregularities located (i) at the base of the F layer, overhead and out to a ground distance of 334 km; and (ii) embedded in the F layer at a ground distance of 500 km. The parabolic F layer used for this calculation has $h_p = 328$ km, $h_m = 450$ km, and $f_oF = 8.4$ Mc/s.)

for irregularities at various heights with respect to the F layer and at certain ground distances east and/or west of an ionosonde on the magnetic equator. It will now be shown how these individual 'ionograms' can be regarded as 'building blocks' in the interpretation and synthesis of spread- F configurations observed on equatorial ionograms.

SIMULATION OF 'EQUATORIAL SPREAD-F' CONFIGURATIONS

Procedure. Configurations of spread echoes regarded here as the superposition of the intensities arising from individual scatterers. 'Ionogram' corresponding to one of these

scatterers will usually be a multiple-valued function of the exploring frequency. If the scattering process is assumed to be weak (i.e., linear) the resultant ionogram produced by groups of irregularities can be synthesized by combining the ionograms associated with the individual scattering centers.

The effect of the intervening ionosphere in defining possible ray paths has been studied, and 'ionograms' have been produced, each corresponding to a single scattering center. It now remains to interpret certain observed equatorial ionograms on the present model.

The simulation of a given ionogram proceeds as follows: First, an $N(h)$ distribution corresponding to the regular F trace of the ionogram is estimated. (Usually, for nighttime Huancayo ionograms, a simple parabolic distribution of electron density is a good approximation.) Next, from certain characteristic features of the spread- F configurations, an array of scattering irregularities can be unambiguously determined such that, upon superposition of their associated scatter traces, agreement with the observed spread is obtained. This procedure thus permits the measurement of various parameters of the patches of thin, elongated irregularities giving rise to the spread- F echoes.

Examples of ionogram synthesis. Three spread- F configurations frequently encountered on equatorial ionograms (Figs. 1, 3, and 4) are recognized as being similar to the theoretical possibilities of Figure 10. Sometimes several of these configurations appear concurrently, as in Figure 11a. In this case, the rectangular spread configuration (the prototype of 'equatorial spread F ') occurs at the lower frequencies, and another form of spread F (attributed to embedded irregularities) occurs in the high-frequency part of the same ionogram.

The simulation (to a first approximation) of the ionogram of Figure 11a proceeds as outlined above, resulting in the scatter traces plotted in Figure 11b. The rectangular part of the simulated ionogram is constructed here of a patch of four irregularities located just at the base of the F layer ($h_o = 328$ km). The first of these irregularities is taken to be overhead, and the others are assumed to extend east or west of the ionosonde progressively outward to a ground distance of 300 km. The triangular part of the simulated ionogram has been obtained by assuming an

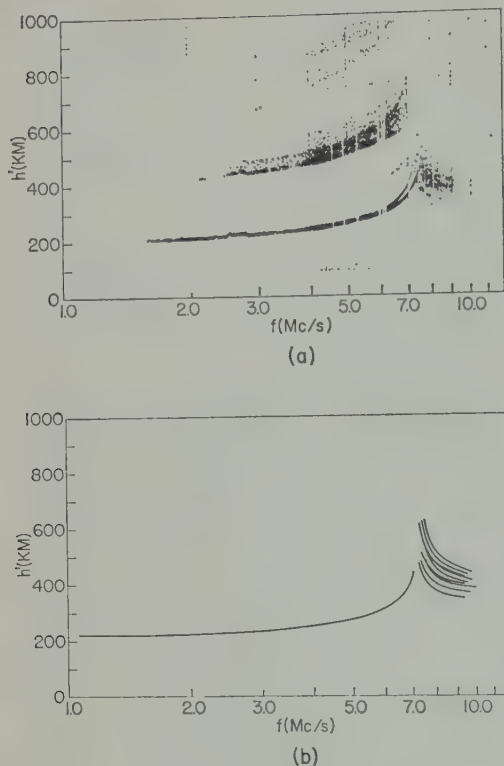


Fig. 12. (a) An ionogram at Huancayo in which a configuration called 'feathers' occurs. Recorded at 0100 EST on October 15, 1960. (b) Simulation of this ionogram, utilizing scattering irregularities from 5 to 45 km above the peak of the F layer and extending from directly overhead out to a ground distance of 100 km. (The parabolic F layer in this case has $h_0 = 215$ km, $h_m = 305$ km, and $f_oF = 7.3$ Mc/s.)

irregularity embedded in the F layer at a height of 334 km and at a ground distance of 500 km. (The high-frequency cutoff of each scatter trace has been introduced arbitrarily.)

The ionogram of Figure 12a is an example of the third distinctive spread- F configuration attributable to thin, field-aligned irregularities. (The descriptive term 'feathers' characterizes its form rather well.) In this case, the irregularities are embedded in the F layer at or above the level of maximum electron density. Thus, on the present interpretation, scattering irregularities are actually being observed above the normal 'ceiling' of the ionosonde (h_m). The synthesis (Fig. 12b) of this configuration is based on the assumption of a patch of irregularities located

from directly overhead out to a ground distance of 100 km, and extending in height between 5 and 350 km (i.e., having a thickness of 40 km) where $h_0 = 215$ km and $h_m = 305$ km. The simulation in this case is especially sensitive to the heights of the irregularities.

The ionogram of Figure 13a shows a more complex triangular configuration than that of Figure 11a. The complexity is interpreted as due to the horizontal extent of the patch of irregularities. The synthesis (Fig. 13b) of this ionogram is based on four irregularities at height of 227 km and at ground distances of 250, 300, 350, and 400 km east or west of the ion

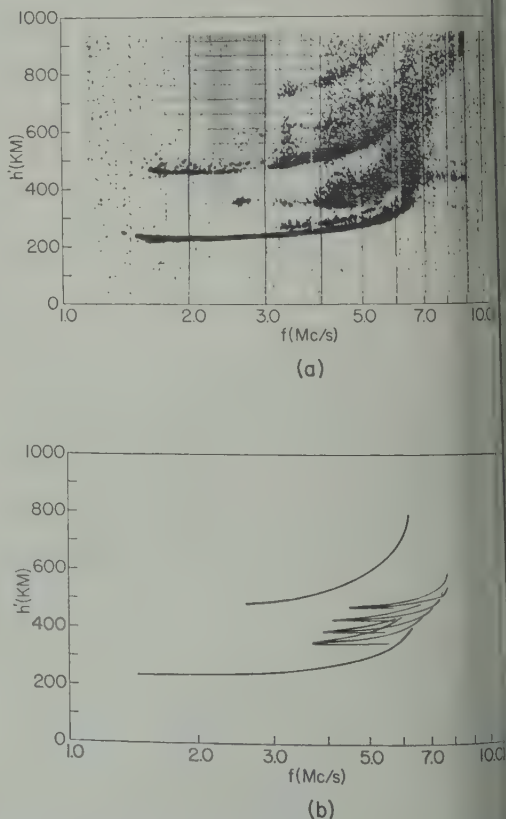


Fig. 13. (a) An ionogram exhibiting a configuration similar to that of Figure 3. Recorded at Huancayo at 0015 EST on June 17, 1960. (b) Simulation of this ionogram, with four scattering irregularities situated 7 km above the base of the F layer and at ground distances of 250, 300, 350, and 400 km. (The parabolic F layer in this case has $h_0 = 227$ km, $h_m = 320$ km, and $f_oF = 6.8$ Mc/s.) The location of the 'triple points' is a linear function of the frequency.

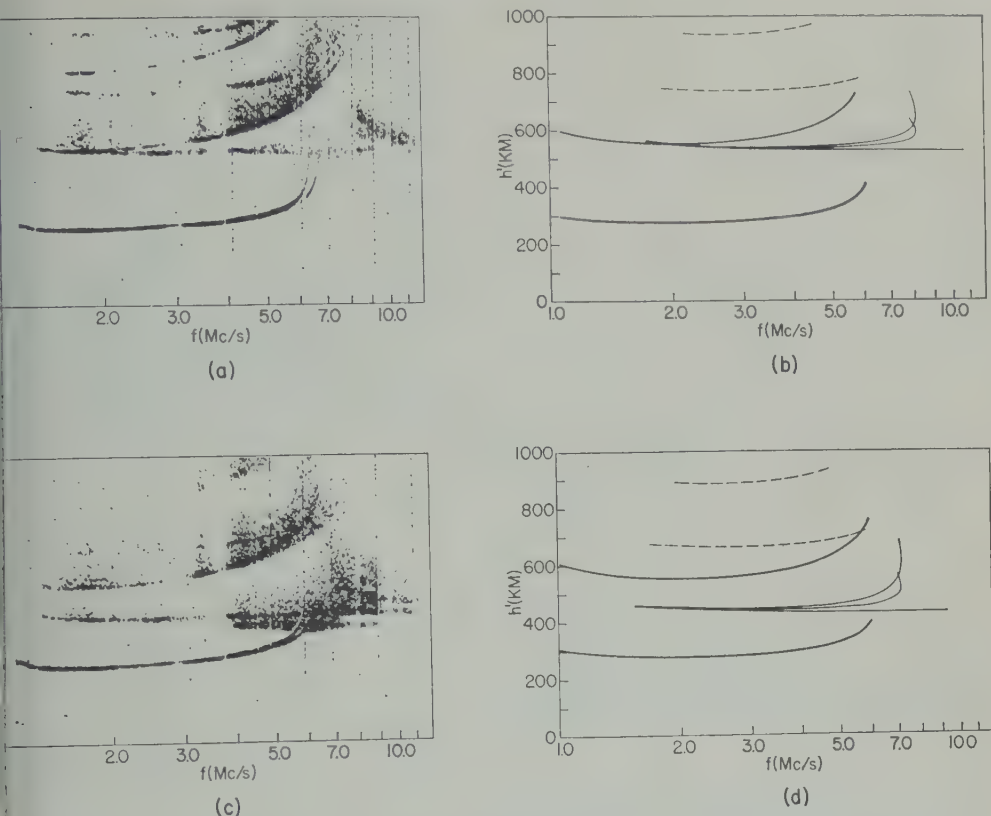


Fig. 14. (a) Huancayo ionogram recorded at 0258 EST on June 23, 1960. (b) Simulation of this ionogram, utilizing an irregularity at a height of 261 km (slightly above the base of the F layer) and at a ground distance of 460 km. (c) Huancayo ionogram recorded at 0315 EST on June 23, 1960, 17 minutes later than (a). (d) Simulation of this ionogram, utilizing an irregularity at a height of 261 km and at a ground distance of 340 km. The electron-density profile used for these calculations was composed of (i) a rectangular E layer (to account for retardation at the lower frequencies) extending from 100 km to the base of the F layer, with $f_oE = 0.62$ Mc/s, plus (ii) a parabolic F layer with $h_o = 260$ km, $h_m = 330$ km, $f_oF = 6.2$ Mc/s.

(The base of the F layer in this case is at a height of 220 km.) From this procedure the uniqueness of the synthesis is established to have an east-west extent of about 150 km.

Interpretations and decompositions. The syntheses in the preceding examples do not uniquely establish, without additional information, whether the irregularities are located east or west of the station. Furthermore, the synthesis of a spread-F configuration is not necessarily unique because different interpretations and decompositions of a spread-F configuration might be possible on the basis of a different scattering model. However, the following successful synthesis of a continuous sequence of ionograms representing

the drift of a patch of irregularities (assumed to travel at a fixed height) is a convincing demonstration of the uniqueness of the present model.

Two ionograms from such a sequence are shown in Figures 14a and 14c, taken 17 minutes apart. Their partial synthesis is indicated in the adjacent diagrams of Figures 14b and 14d, where a single irregularity has been located just above the base of the F layer, at a height of 261 km, and at a ground distance of 460 and 340 km, respectively. In these diagrams, 'multiples' are apparently associated with the scatter traces. These additional traces arise from propagation paths involving reflection by the F layer with an intermediate ground reflection before arriving

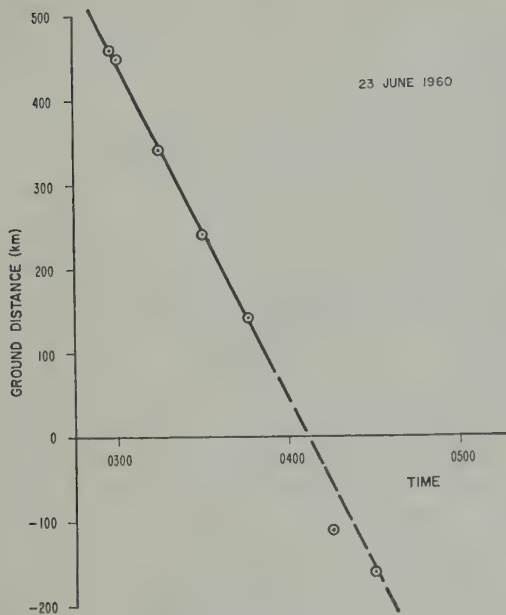


Fig. 15. The ground distance of a patch of irregularities plotted against the local time of the Huancayo ionogram used in the calculation, for a sequence of ionograms. The first and third points result from Figure 14.

at the scatterer. Their equivalent range may be computed by the above techniques. In Figures 14b and 14d the lower dashed trace corresponds to one intermediate ground reflection during the round trip, and the upper dashed trace results from an intermediate ground reflection both before and after scattering.

Figure 15 is a summary of the ground-distance measurements for the entire sequence studied, including the points at 460 and 340 km that correspond to the ionograms of Figure 14. The ground distance of the patch decreases, then increases. The interpretation of this variation is that the patch has passed overhead. The linearity of this plot implies that the assumption that the patch of irregularities drifts at constant height is a reasonable one. The drift-speed component in the east-west vertical plane measured from Figure 15 is 113 meters/sec. (Such drift speeds are for the motion of *patches* of irregularities, and may not pertain to the *individual* irregularities.)

Although this particular measurement of drift speed does not establish whether the drift is from west to east or vice versa, such a sensing

of direction is possible if further data are available. Suitable data for this purpose were deduced for the period December 1957 through November 1958 from transequatorial scattering measurements between Antofagasta, Chile ($23^{\circ}44'70''15''W$), and Guayaquil, Ecuador ($2^{\circ}36'80''24''W$) [Cohen and Bowles, 1961]. The field aligned irregularities visible from Huancayo, owing to the orthogonality requirement, are only those along the dashed magnetic isoclines slightly to the north (Fig. 16), while the scattering volume (defined by the antenna-beam intersection) in which irregularities may be detected over the oblique path is located to the southwest of Huancayo. Because of this asymmetry

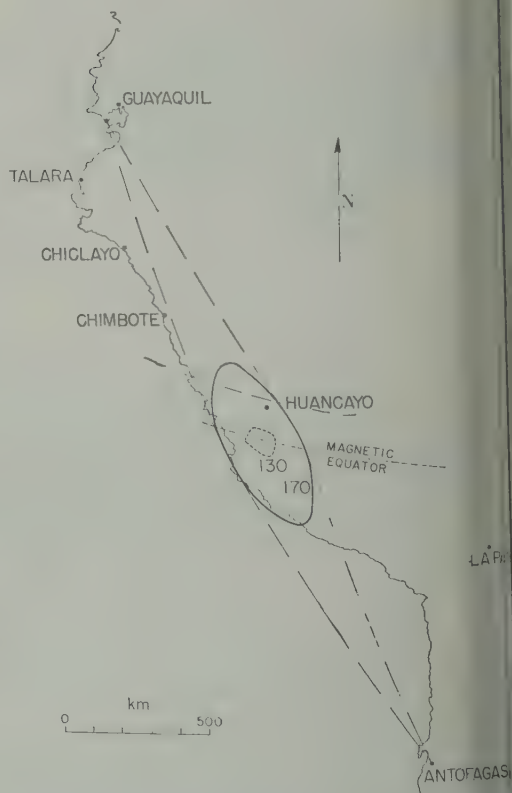
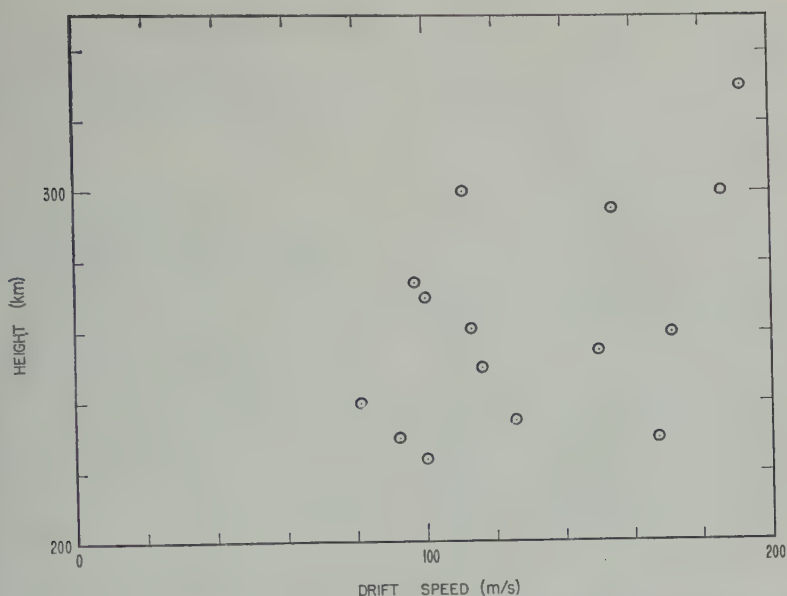


Fig. 16. Map of the west coastal area of South America with the approximate loci of antenna-beam intersections (between Antofagasta and Guayaquil) computed for the heights indicated (130 and 170 km). The dashed line just north of Huancayo represents the locus of points at which orthogonality to magnetic-field-aligned irregularities at a height of 300 km is achieved from Huancayo. (La Paz, Bolivia; Talara, Peru; Chiclayo, Peru; and Chimbote, Peru, are the locations of other ionosondes.)



17. Drift-speed determinations vs. the height of the patch of irregularities for each case.

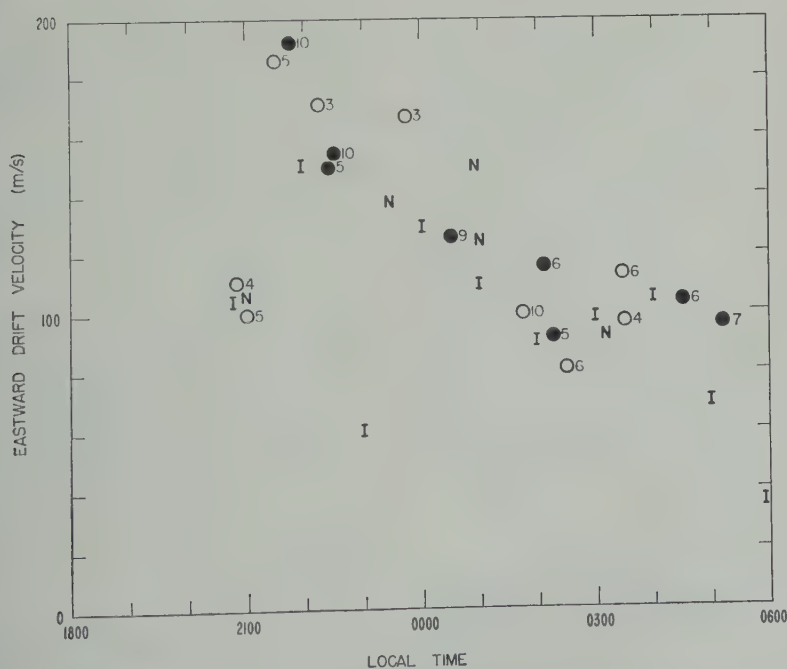


Fig. 18. Various determinations of drift velocities, as follows: (i) For patches of 'equatorial spread F' ' irregularities observed at Huancayo, Peru, the filled circles represent determinations for which sensing of direction was possible; the open circles represent cases which could not be sensed. The number adjacent to each circle corresponds to the month of observation. (ii) For patches of equatorial spread F' irregularities observed at Natal, Brazil, drift-speed determinations are represented by the symbol N . (iii) For the F -layer motions observed at Ibadan, Nigeria, the drift velocities measured by Skinner, Hope, and Wright [1958] have been plotted using the symbol I .

patches of irregularities drifting in the vertical plane of visibility from Huancayo would be detected most sensitively on the oblique path somewhat earlier or later than their overhead passage at Huancayo, depending on whether their drift is from west to east or vice versa. Since the observed drift velocities are between 100 and 200 meters/sec, this east-west displacement corresponds to a relative time delay of 15 to 30 minutes between the maxima of signal enhancements on the scatter path and the transit above Huancayo.

By means of the techniques described above, a series of drift-speed measurements has been obtained, and, whenever possible, the sense of the motion has also been determined. The possibility that drift speed might be a function of height has been examined in Figure 17, with an inconclusive result. On the other hand, a nocturnal trend of these speeds is apparent (Fig. 18), with speeds decreasing from about 200 meters/sec after 2100 local time to around 100 meters/sec at 0400. There was, however, a tendency for the occurrence of lower drift speeds prior to 2100. When a sense determination could be accomplished, the drift velocity was established to be *eastward*. (Drift-speed determinations that could not be sensed have also been included as eastward drift velocities in Fig. 18.) In addition, five drift speeds have thus far been determined for Natal, Brazil (5.3°S, 35.1°W; magnetic dip, 2°S), by the procedures outlined above. These measurements are plotted as eastward velocities in Figure 18, using the symbol *N*.

Drift velocities of the equatorial *F* layer have been measured at Ibadan, Nigeria, by Skinner, Hope, and Wright [1958], by means of spaced-antenna techniques. Their results are reproduced in Figure 18, using the symbol *I*. Further drift measurements are being made at Ibadan [Lyon, Skinner, and Wright, 1960].

The close agreement between the drift velocities determined at such widely spaced locations is remarkable. Furthermore, the drift velocities measured along the magnetic equator in South America were those of extensive patches of irregularities, whereas the drift velocities measured at Ibadan were those of diffraction patterns produced at the reflection level.

Knecht says that he interprets as a drift velocity the delay in onset of spread *F* at Talara, Chiclayo, Chimbote, and Huancayo (Figure 16),

Assuming the drift to be along the magnetic isoclines, his statistical analysis indicates eastward velocities consistent with the measurements presented in Figure 18 [Knecht, 1960; Hirsch, Knecht, 1962].

CONCLUSIONS

1. Certain spread-*F* configurations observed at the magnetic equator arise from scattering in the vertical, east-west plane from thin, field-aligned irregularities. Their differences in appearance depend on the positions of the irregularities in the *F* layer, and result solely from refractive effects imposed by the intervening ionosphere. Accordingly, the 'equatorial spread *F*' category is broadened to include all the ionogram configurations associated with such irregularities. (Further such revisions of terminology, based on the physical processes involved rather than on the appearances of ionograms, may be anticipated.)

2. The 'ionograms' corresponding to individual scattering centers constitute the 'building blocks' from which, by superposition, the spread configurations actually occurring can be synthesized. In this process, various parameters of patches of irregularities can be determined.

3. The patches of 'equatorial spread *F*' irregularities can occur at heights as much as 100 km below the base of the equatorial *F* layer and 50 km or more above the level of maximum electron density.

4. The extent of these patches in the east-west plane may be as great as 150 to 300 km, and they can be detected and localized in this plane.

5. The patches of irregularities have a thickness ranging between 10 and 50 km.

6. The drift velocity of such patches is eastward, gradually decreasing from 200 to 100 meters/sec after 2100 hours local time. However, during the preceding evening hours, drifts at lower speeds can occur.

APPENDIX

Transformation Relationships for 'Backscatter' Rays

The equivalent range from an ionosonde to scattering irregularity is

$$h' = \int_{\text{ray}} \mu' ds$$

is the group refractive index and ds is
ent of path length.
magnetic field is neglected, the dispersion
has the form

$$\mu^2(f) = 1 - f_N^2/f^2 \quad (2)$$

is the phase refractive index, f_N is the
asma frequency, and f is the exploring
y. It can be shown using (2) that

$$\mu\mu' = 1 \quad (3)$$

figure 5, it can be seen that

$$ds = dz/\cos \phi \quad (4)$$

tion of (3) and (4) into (1) gives

$$dz/\mu \cos \phi$$

ny

$$= \int_{\text{ray}} |dz| (\mu^2 - \mu^2 \sin^2 \phi)^{-1/2} \quad (5)$$

's law,

$$\mu \sin \phi = \text{constant} = \mu_A \quad (6)$$

μ_A is the index of refraction at the apex
ay path; i.e., from (2),

$$\mu_A^2 = 1 - f_A^2/f^2 \quad (7)$$

tion of (2) and (7) into (5) yields

$$h' = f \int_{\text{ray}} |dz| (f_A^2 - f_N^2)^{-1/2} \quad (8)$$

egral of (8) is constant for rays that
the same height interval as long as
rays have reflection points at the same
(the height characterized by the plasma
ey f_A). Thus, for such a family of ray
here is the relationship

$$h'/f = \text{constant} \quad (9)$$

e subscript V refer to a ray at vertical
ce, and the subscript R refer to a corre-
g oblique ray, i.e., two rays for which
lies. Then, from (9)

$$h'_V/f_V = h'_R/f_R \quad (10)$$

her useful transformation may be derived
egration in terms of the horizontal co-
e. From Figure 5,

$$ds = dx/\sin \phi \quad (11)$$

so that (1) may be written, in view of equations
3 and 6,

$$h' = \int_{\text{ray}} dx/\mu \sin \phi$$

$$= \int_{\text{ray}} dx/\mu_A = R/\mu_A \quad (12)$$

where R is the ground distance of the irregularity.
Substitution of $\mu_A = R/h'$ from (12) into (7)
yields

$$h'^2 = (f_A^2/f^2)h'^2 + R^2 \quad (13)$$

At oblique incidence, $h' \rightarrow h'_R$, and $f \rightarrow f_R$.
Also, from (7), since normal-incidence reflection
occurs where $\mu^2 = 0$,

$$f_A = f_V \quad (14)$$

so that (13) becomes

$$h'^2_R = (f_V^2/f_R^2)h'^2_R + R^2 \quad (15)$$

Using (10), this equation can be written

$$h'^2_R = h'^2_V + R^2 = h'^2_V(1 + R^2/h'^2_V) \quad (16)$$

$$\text{or } h'_R = h'_V(1 + R^2/h'^2_V)^{1/2} \quad (16')$$

Alternatively, (16') may be expressed, following
(10), as

$$f_R = f_V(1 + R^2/h'^2_V)^{1/2} \quad (17)$$

Equations 16' and 17 are recognized as com-
prising the transformation (2) utilized in the text.

Acknowledgments. The collaboration of K. L.
Bowles, M. L. V. Pitteway, T. E. VanZandt, J. R.
Winkelman, and J. T. Brown, Jr., is gratefully
acknowledged.

REFERENCES

- Bates, Howard F., Direct HF backscatter from the
 F region, *J. Geophys. Research*, **65**, 1993-2002,
1960.
Booker, H. G., A local reduction of F -region
ionization due to missile transit, *J. Geophys.*
Research, **66**, 1073-1079, 1961.
Cohen, Robert, and Kenneth L. Bowles, On the
nature of equatorial spread F , *J. Geophys.*
Research, **66**, 1081-1106, 1961.
Dieminger, W., The scattering of radio waves,
Proc. Phys. Soc. London, **64**, B, 142-158, 1951.

- Eckersley, T. L., Irregular ionic clouds in the *E* layer of the ionosphere, *Nature*, 140, 846-847, 1937.
- Eckersley, T. L., Analysis of the effect of scattering in radio transmission, *J. Inst. Elec. Engrs., London*, 86, 548-563, 1940.
- Gassmann, George J., Measurements of irregularities and drifts in the arctic ionosphere using airborne techniques, *Polar Atmosphere Symposium*, part II, Ionospheric Section, Pergamon Press, London, 44-51, 1957.
- Hirsh, A. J., and R. W. Knecht, On the mean temporal variations of electron density at a fixed height in the *F* region, *J. Geophys. Research*, 67, in press, 1962.
- Knecht, R. W., Possibility of detecting ionospheric drifts from the occurrence of spread *F* echoes at low latitudes, *Nature*, 187, 927, 1960.
- Lyon, A. J., N. J. Skinner, and R. W. Wright, Preliminary results of horizontal drift measurements in the *F* layer near the magnetic equator, *Some Ionospheric Results Obtained during the International Geophysical Year*, Elsevier Publishing Company, Amsterdam, 333-337, 1960.
- McNicol, R. W. E., H. C. Webster, and G. G. Bowman, A study of spread-*F* ionospheric echoes at night at Brisbane, 1, Range spreading (experimental), *Australian J. Phys.*, 9, 247-271, 1956.
- Martyn, D. F., The propagation of medium radio waves in the ionosphere, *Proc. Phys. Soc. London*, 323-339, 1935.
- Peterson, Allen M., The mechanism of *F*-layer propagated back-scatter echoes, *J. Geophys. Research*, 56, 221-237, 1951.
- Pitteway, M. L. V., and Robert Cohen, A waveguide interpretation of 'temperate-latitude spread *F*' on equatorial ionograms, *J. Geophys. Research*, 66, 3141-3156, 1961.
- Ratcliffe, J. A., *The Magneto-Ionic Theory and Applications to the Ionosphere*, Cambridge University Press, 206 pp., 1959.
- Renau, Jacques, Theory of spread *F* based on aspect-sensitive backscattered echoes, *J. Geophys. Research*, 65, 2269-2277, 1960.
- Skinner, N. J., J. Hope, and R. W. Wright, Horizontal drift measurements in the ionosphere near the equator, *Nature*, 182, 1363-1365, 1958.

(Manuscript received July 17, 1961.)

A Waveguide Interpretation of 'Temperate-Latitude Spread F ' on Equatorial Ionograms

M. L. V. PITTEWAY¹ AND ROBERT COHEN

*Central Radio Propagation Laboratory, National Bureau of Standards
Boulder, Colorado*

Abstract. A waveguide model is presented for propagation of radio waves along elongated irregularities aligned parallel to the earth's magnetic field in the equatorial ionosphere. This theory is applied to analyze the frequency spread ionograms often observed during the equatorial night, particular attention being devoted to the detailed striations of the spread; similar spread- F striations on arctic ionograms are explained by a corresponding waveguide theory. The waveguide theory is consistent with direction-finding evidence that the category of spread F considered in this paper is caused by back-scatter in the north-south plane from irregularities that *do not* support north-south forward scatter at 50 Mc/s. These echoes are in contrast to those arriving in the east-west plane from the 'equatorial spread F ' irregularities considered elsewhere, which *do* support north-south forward scatter at 50 Mc/s. This paper is concerned with propagation in an irregular ionosphere, but not with the formation of the irregularities.

1. INTRODUCTION

A standard radio technique for exploring the ionosphere, the time taken by a short radio pulse to reach, be reflected by, and return from the ionosphere is exhibited on an oscilloscope as a function of increasing frequency. Two traces are normally observed (apart from multiple reflections between the ionosphere and the ground), the ordinary and extraordinary ray paths; as the frequency is increased, the pulses are reflected from higher inside the ionosphere, and increasingly large time delays are produced at the critical frequencies for penetration of the ordinary E or F layers.

Sometimes at night the echoes from the F layer are spread in delay times, and this spread is too great to be explained solely by group velocity dispersion in a horizontally stratified ionosphere. It is concluded that this 'spread F ' is caused by back-scattering from irregularities in the ionosphere. The basic categories of spread F have been distinguished on equatorial ionograms [Cohen and Bowles, 1961]. One category, known as 'equatorial spread F ', is attributed to echoes from field-aligned irregularities which arrive in the east-west plane passing through the ionosonde [Cohen and Cohen, 1961]; the other, known as 'temperate-latitude spread F ', is the subject of

this paper. (The terms 'equatorial spread F ' and 'temperate-latitude spread F ' refer in this paper only to ionogram configurations observed near the magnetic equator. The former is characteristic of equatorial ionograms; the latter is similar in appearance to the spread F sometimes observed on temperate-latitude ionograms. 'North,' 'east,' 'south,' and 'west' always refer to magnetic directions.)

The standard ray-theory treatment of propagation in an irregular ionosphere is not strictly valid near heights where the radio wave is reflected. The spread- F mechanism described in this paper occurs under precisely these conditions and it is therefore necessary to use the differential wave equations instead. These equations can be solved explicitly for *weak* irregularities by means of the Born approximation [Pitteway, 1958, 1960].

Booker and Pitteway [1961] consider from a waveguide standpoint the *strong* scatter by irregularities elongated in the direction of the earth's magnetic field, in order to explain the formation of spread F on arctic ionograms. Their work is adapted here to explain the similar 'temperate-latitude spread F ' on equatorial ionograms.

At Huancayo, Peru (12°3'S, 75°20'W; magnetic dip, 2°N), field-aligned irregularities are elongated almost horizontally in a north-south direction. In considering spread F caused by waveguide propagation along these irregularities, no assumption is made about their cross section;

¹ Present address: DSIR Radio Research Station, Slough, Park, Slough, Buckinghamshire, England.

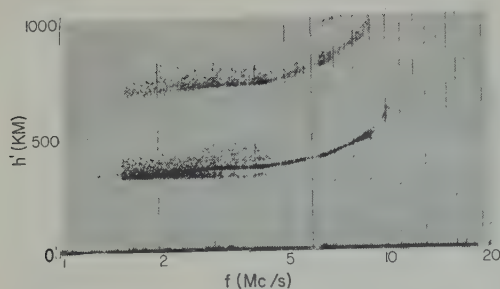


Fig. 1a. A rectangular 'equatorial spread F ' configuration at Huancayo due to irregularities situated low in the F layer; 2100 EST, October 29, 1960.

they might have cylindrical symmetry, or they might be field-aligned planes analogous to the E -layer irregularities believed to be associated with the equatorial electrojet [Bowles and Cohen, 1961].

A summary of the spread- F effects at the equator is given in section 2 of this paper, together with the earlier explanation of 'equatorial spread F .' A simple waveguide theory which explains the 'temperate-latitude spread F ' is developed in section 3, and is used in section 4 for the detailed interpretation of ionograms. Experimental evidence is presented in section 5 to show that the 'temperate-latitude spread F ' at Huancayo is indeed due to propagation in north-south waveguides, as suggested. Further theory in section 6 provides a quantitative basis for future work.

2. SPREAD- F IRREGULARITIES IN THE EQUATORIAL IONOSPHERE

Scattered echoes from F -region irregularities frequently appear on equatorial ionograms, and on occasion are sufficiently intense to obscure the regular F trace. Usually the spread- F echoes occur during the hours when the F region is in darkness, although during years of low sunspot activity there are instances of spread F continuing into the morning hours. The spread F observed on equatorial ionograms occurs in various configurations, which can be classified by their appearance.

A more fundamental classification of the spread- F echoes can be achieved by a study of the ionospheric irregularities giving rise to them. Such a study has been conducted by means of oblique incidence scattering experiments at 50 Mc/s over an approximately north-south path.

One category of irregularities, often associated with a rectangular spread- F configuration in the Huancayo ionograms, supports oblique scatter propagation; the other category, associated with 'temperate-latitude spread F ,' does not [Cohen and Bowles, 1961]. The term 'equatorial spread F ' was originally reserved for the configurations of rectangular shape, such as that shown in Figure 1. It has since been recognized [Calvert and Cohen, 1961] that there are other strikingly different configurations, which are associated with oblique scatter propagation, and should therefore be included in the 'equatorial spread F ' category. Examples of two such configurations are shown in Figures 2 and 3.

This 'equatorial spread F ' has been explained as echoes (from thin, field-aligned irregularities at least 1 km long) which arrive in the vertical east-west plane passing through the ionosphere. The oblique scatter at 50 Mc/s indicates the presence of irregularities with a vertical dimension less than about 10 meters. These irregularities occur in patches throughout the F region. The effect upon the ionograms is determined by whether they are (a) located at or below the base of the F layer, (b) embedded in the F layer at or below its electron density maximum, or (c) embedded in the F layer above its electron density maximum.

Case (a) gives the rectangular configuration of Figure 1; group retardation en route to the scattering patches of cases (b) and (c) leads to the configurations of Figures 2 and 3, respectively. Subject to a single-scatter approximation the patches of irregularities act as a superposition of individual scatterers. The geometrical

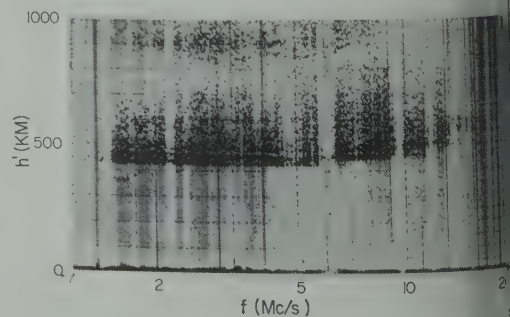
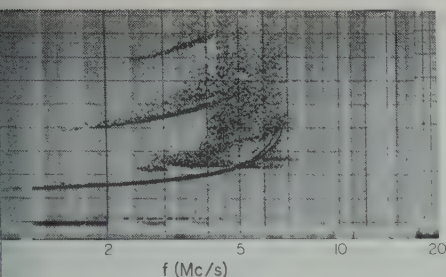


Fig. 1b. A stronger manifestation of 'equatorial spread F ' at Huancayo, with some of the irregularities higher in the F layer; 2016 EST, April 9, 1961. (The upper configurations result from ground reflections.)



2. 'Equatorial spread F ' at Huancayo due to irregularities embedded in the F layer below the height of maximum electron density; 0215 EST, 10, 1958.

uttering by a single irregularity at the different ionospheric heights is sketched in Figure 4, together with the corresponding 'ams.'

The second category, associated with 'temperate-latitude spread F ,' does not support scatter at 50 Mc/s, and so must consist of irregularities with vertical scale size less than 10 meters. It was suggested [Cohen and Ingham, 1961] that these irregularities might be of all sizes greater than even the ionogram wavelengths, in which case those near the bottom level would be most effective as scatterers. The present paper is a detailed investigation of the 'temperate-latitude spread F ' phenomenon.

Main features of the 'temperate-latitude spread F ' on Huancayo ionograms suggest the waveguide model discussed here. Most significant is the presence of marked striations, as in the traces of Figure 5, with slope much steeper than that of the upper trace of Figure 4b. The variation of the 'temperate-latitude spread F ' is interesting; the higher-frequency part of the ionogram often is gradually filled in (Fig. 6) but there appears to be a disturbance propagating downward through the ionosphere, starting from an electron-density maximum and leaving irregularities in its wake. The time sequences of Figure 5 show that as time progresses the striations are observed at lower and lower frequencies. In contrast to the mechanism shown in Figure 4c, it turns out (section 4) that there is no waveguide propagation if the irregularities are situated above the electron-density maximum of the F layer.

There is yet another class of spread- F echoes observed on equatorial ionograms, known as

'satellite traces,' such as in Figure 7. These are usually the first manifestation of spread F after sunset, and are attributed to large-scale distortions of the isoionic contours and consequent oblique reflections. These satellites frequently precede an evening display of 'equatorial spread F ,' and 'temperate-latitude spread F ' often appears by midnight.

3. A SIMPLIFIED WAVEGUIDE MODEL

In this paper, the 'temperate-latitude spread F ' on equatorial ionograms, including in particular the detailed striations, is interpreted as arising from waveguide propagation along irregularities elongated along the earth's magnetic field. In contrast with the thin irregularities responsible for the 'equatorial spread F ' of section 2, waveguide propagation requires thick, large-scale irregularities which are effective only near the level of reflection (Fig. 8).

This section is devoted to a simplified wave-theory model of guided propagation. To avoid mathematical complexity, east-west variations in electron density are ignored, and so the problem reduces to two dimensions, z (vertically upward), and x (north-south).

A plane, monochromatic wave of frequency f is considered incident onto the ionosphere at an angle θ to the vertical in the magnetic meridian, where $C = \cos \theta$; the unit of length is defined as the free-space wavelength, λ_0 , and a time factor, $\exp(2\pi i f t)$, is omitted throughout this section. The effects of the earth's magnetic field and of collisional damping are ignored, so that the ionospheric refractive index is defined by the standard parameter $X = f_N^2/f^2$, where f_N^2 is proportional to the electron density. For

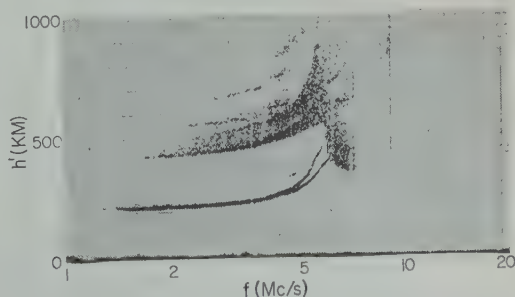


Fig. 3. 'Equatorial spread F ' at Huancayo due to irregularities embedded in the F layer above the height of maximum electron density; 0414 EST, April 28, 1960.

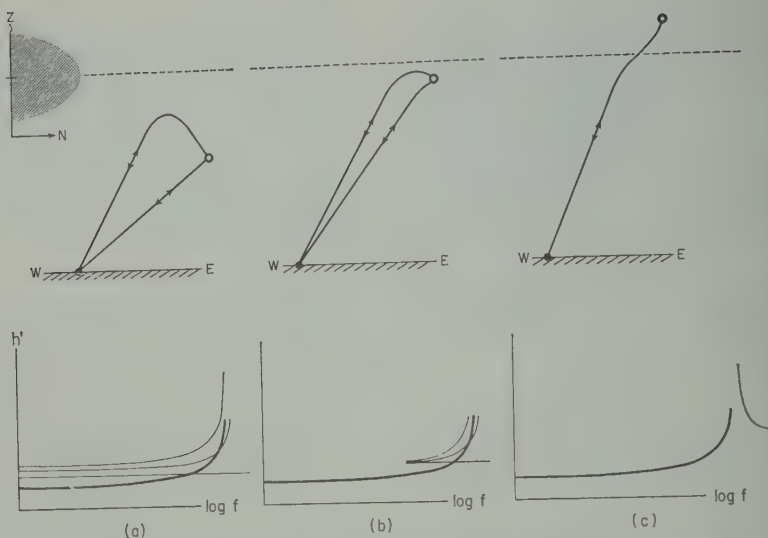


Fig. 4. Ray paths and 'ionogram' for an irregularity (a) located at or below the F layer; (b) embedded in the F layer at or below the height of maximum electron density; or (c) embedded in the F -layer above the height of maximum electron density.

polarization such that the electric wave-field vector E is parallel to the y axis, there are no induced space charges, and Maxwell's equations, together with the constitutive relations for the ionosphere, reduce to

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial z^2} + 4\pi^2(1 - X)E = 0 \quad (1)$$

For the model of this section, the ionospheric parameter $X(x, z)$ is composed of two parts, representing a smooth horizontally stratified ionosphere plus elongated irregularities which vary slowly in the x direction. First, suppose that X is independent of x ; then the x variation

which is modified to $X = X_0 - \Delta$ for $-a < z < a$; this 'waveguide irregularity' will support a trapped mode of propagation if $\Delta > 0$. However, trapped modes are possible only for certain discrete values of the angle parameters in (2), which must satisfy the boundary conditions that E and dE/dz be continuous at $z = \pm a$. Clearly

$$X_0 - \Delta < C^2 < X_0$$

because $X_0 - \Delta$ must be underdense, and X_0 overdense.

The modes can be classified as symmetric and antisymmetric. Solving (2) gives:

		Symmetric	Antisymmetric
$z > a$	$X = X_0$	$E = A \exp(-2\pi Qz)$	$E = A \exp(-2\pi Qz)$
$-a < z < a$	$X = X_0 - \Delta$	$E = B \cos(2\pi qz)$	$E = B \sin(2\pi qz)$
$z < -a$	$X = X_0$	$E = A \exp(2\pi Qz)$	$E = -A \exp(2\pi Qz)$

of E is described by a factor $\exp(-2\pi i Sx)$, where $S = \sin \theta$, and omitting this factor, E satisfies

$$\frac{d^2 E}{dz^2} + 4\pi^2(C^2 - X)E = 0 \quad (2)$$

A simple square well variation for X is shown in Figure 9; the smooth background ionosphere is represented by a constant value $X = X_0$,

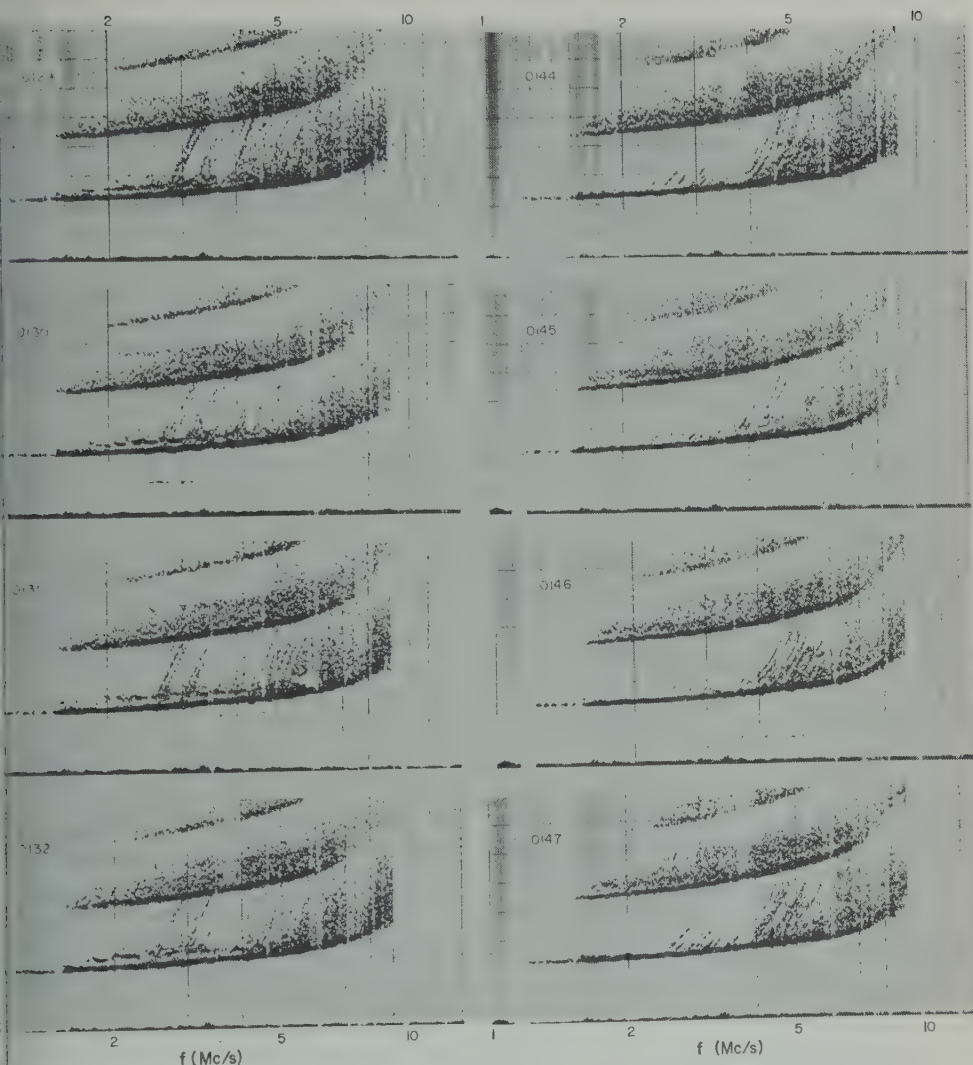
and equating E , dE/dz at $z = a$: $Q = q \tan(2\pi qa)$

$$Q = -q \cot(2\pi qa)$$

where

$$Q = \sqrt{X_0 - C^2} \quad q = \sqrt{\Delta - X_0 + C^2}$$

Both Q and q are real and positive. (Note that if $a^2 \Delta < 1/16$, only one mode is possible, and it will be symmetric; further, if $a^2 \Delta \ll 1/16$,



5. Two sequences of Huancayo ionograms for March 12, 1960, showing the downward motion of striations. *Left*, 0129–0132 EST; *right*, 0144–0147 EST.

aq^2 , and if $a^2 \Delta \ll 1/16\pi^2$, $Q \approx 2\pi a\Delta$.) Suppose that the lower wall of the waveguide ($z < -a$) has a finite thickness b , and now the level $z = -(a+b)$ the ionosphere again becomes underdense. Energy will be trapped by the barrier $-(a+b) < z < -a$ to a leakage factor $\exp(-2\pi Qb)$, which is appreciable (say, $Qb \leq 1/2$, $\exp(-\pi) \geq \exp(-2\pi) = 0.043$), there will be trapped modes with discrete eigenvalues for $Qb \geq 1/2$, $b^2(X_0 - C^2) \geq 1/4$, and since $X_0 - C^2 < \Delta$, we require $b^2 \Delta > 1/4$ for a trapped mode. At 3 Mc/s, $\lambda_0 = 100$ meters;

if Δ is about 0.04, $b > 250$ meters. Similar order-of-magnitude calculations would apply to a more realistic electron-density profile like that of Figure 10, and therefore to trap energy at 3 Mc/s a barrier of increased electron density some 250 meters thick is necessary.

Next, slow north-south (x) variations are introduced into the model. Strictly speaking, (2) should be discarded for the partial differential wave equation 1, but with a slowly varying approximation (2) is retained, bearing in mind that small variations of C are possible inside the ionosphere. As x changes, a waveguide mode

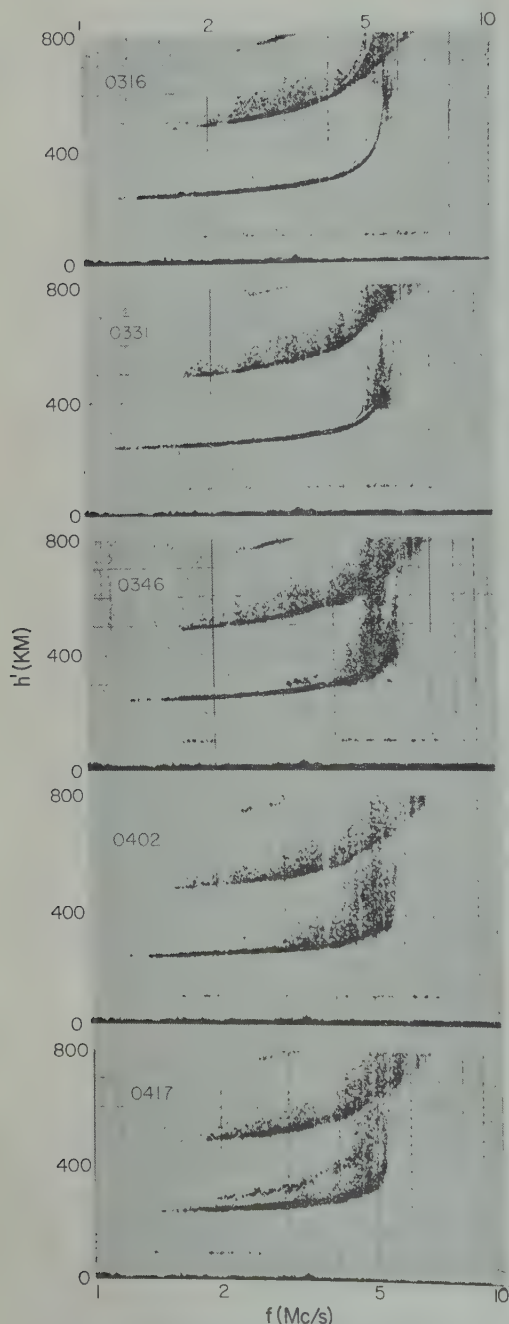


Fig. 6. A sequence of Huancayo ionograms exhibiting the 'cascading' of 'temperate-latitude spread F'; 0316, 0346, 0402, and 0417 EST, May 20, 1960.

can change from leaking to trapped (Fig. 1) producing a small change in C , from that corresponding to the angle of incidence to the discrete value for mode propagation. The energy contained in a wave incident obliquely from below is divided in the leaky region, some of the trapped energy being returned from internal reflection in the waveguide, and the resulting echo is observed at the ground. Then, as the exploring frequency increases, the point of reflection is pushed further back into the waveguide, until penetration of the wall occurs.

The waveguide modes of propagation can only be excited near the levels of regular ionospheric reflection, Figure 12. At vertical incidence, the large 'wavelengths' near the reflection level allow the scattering through 90° required to enter the mode propagation. (In terms of wave theory, the slow x variations are sufficient to couple energy into the waveguides only when $X \approx 1$.) At oblique incidence, energy can be channeled by the elongated irregularities only after regular refraction has bent the ray almost horizontal. Note that internal reflection in the waveguide provides a mechanism for backscatter only, and so it will only be observed with a receiver near the transmitter.

Had the incident wave been polarized with its electric vector in the xz plane, the effect of induced space charges would add another term to the differential wave equation, which would be most effective for a sharp-walled irregularity like that shown in Figure 9. However, there will also be induced space charges for the first polarization when east-west variations are taken into account; in addition, there will be a separation of the traces for the two polarizations because of the effect of the earth's magnetic

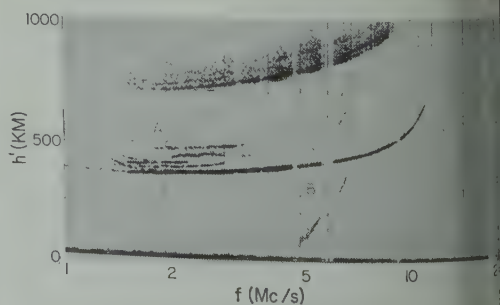


Fig. 7. Satellite traces (A) and high-order multiples (B) of the F trace at Huancayo; 1830 EST, October 30, 1960.



8. Waveguide propagation along large-elongated irregularities situated near the critical reflection level.

Therefore a more complicated mathematical model is indicated.

INTERPRETATION OF THE EQUATORIAL IONOGRAMS

In this section, the waveguide model of section 2 is applied to the 'temperate-latitude spread F ' observed in section 2. The experimental verification of this interpretation is given in section 5. Each individual striation in Figure 5 is interpreted as propagation along a particular north-south aligned 'hole' in the ionosphere. The echo of the waveguide is first seen near the main trace when the exploring wave, of frequency f_1 (Figure 13), reaches the embedded irregularity. This reflection-level phenomenon can never occur below the main trace. At slightly higher frequencies, f_2 and f_3 ($f_3 > f_2 > f_1$), an irregularity at the same height will be effective at a small oblique incidence, and the echo is produced partly by the time delay en route to the reflection height and partly by group velocity dispersion in the waveguide.

As a start, the time delay of group propagation inside the 'waveguides' is ignored, the time of the striations should be determined by ray paths of Figure 13, assuming a fixed reflection height. These shapes are calculated from magnetoionic theory in section 6; departures from this shape might be interpreted as group delays in the waveguides or as large-angle departures of the alignment from the horizontal. For example, the striations in Figure

14 show an interesting flat portion, which might well be caused by a deformed irregularity; the two similar traces are interpreted as the ordinary and extraordinary magnetoionic components, their frequency separation at the points of departure from the main trace being consistent with the equatorial gyrofrequency (section 6).

The 'temperate-latitude spread F ' echoes near the regular penetration frequency are interpreted as a composite of striations, perhaps complicated by additional spreading due to partial penetration of the layer.

Thick irregularities are necessary for waveguide trapping (accounting for their failure to support oblique scatter at 50 Mc/s), in contrast with the thin irregularities required to produce 'equatorial spread F '. On the other hand, the waveguide irregularities cannot be thicker than about 10 km; otherwise, the regular F -layer trace would exhibit more structure than is observed.

Note that, since the waveguide propagation is only effective at the reflection level, irregularities located above the F -layer electron-

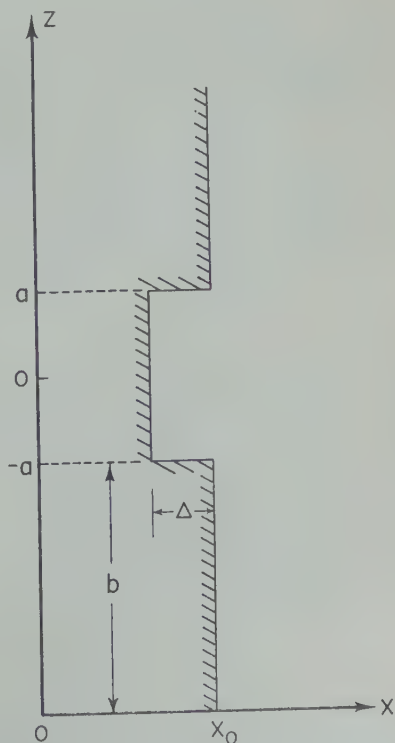


Fig. 9. Simple model for the waveguide analysis.

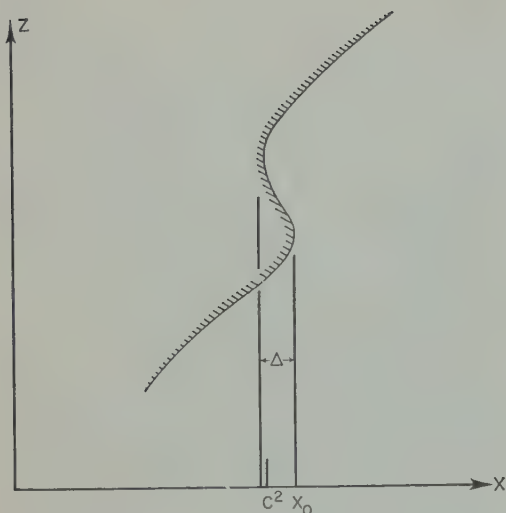


Fig. 10. Electron-density profile for leaky mode propagation.

density maximum cannot produce 'temperate-latitude spread F ,' since regular reflection only occurs below the F -layer maximum.

Inside the ionosphere, there is a tendency for the ray path defining group propagation to deviate from the plane of incidence. The ordinary magnetoionic component turns away from the earth's magnetic field, whereas the extraordinary ray turns toward it [Budden, 1961]. However, the *wave-normal* direction, rather than the *ray* direction, is the important feature in the waveguide model, and so this deviation should not influence the echoes. (Similarly, for the formation of 'equatorial spread F ,' the condition for scattering is that the *wave-normal* be perpendicular to the elongated irregularities.)

Relatively rare 'festoon' echoes are shown in Figure 15; such echoes are shorter lived than the striations, with a duration of about 1 minute.

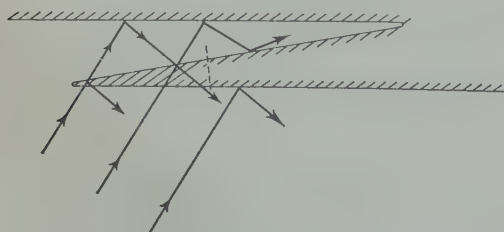


Fig. 11. Wave-theory excitation of mode in a leaky region (compare Fig. 8).

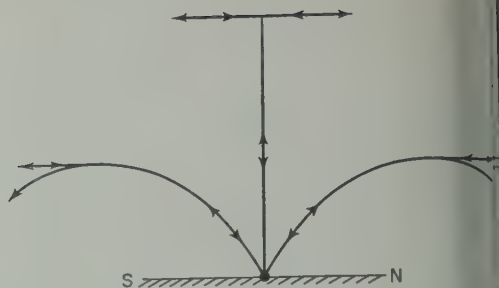


Fig. 12. The excitation of waveguide propagation near the top of the regular ray trajectories.

A suggested propagation model for such echoes is a strong 'hole' or 'waveguide,' the downward trace being accounted for by the initial penetration into this 'hole.'

Comparison with 'arctic spread F .' Since 'temperate-latitude spread F ' was named because of its similarity to the spread- F configuration observed at nonequatorial latitudes, it is interesting to compare and contrast the preceding phenomena with spread F of the arctic variety Figure 16. A waveguide model analogous to that of section 3 has been developed for the arctic case, for which the irregularities are assumed to be vertically elongated [Booker and Pitteway, 1961]. Extending this work further, the detailed striations of Figure 16 are interpreted as echoes resulting from individual modes of propagation in vertical 'holes.' Penetration of the total ionosphere occurs first for the strongest 'hole,' and successive penetrations through other holes build up the composite ionogram of Figure 16. The similarity of 'arctic spread F ' to 'temperate-latitude spread F ' seen at the equator is obvious, both in appearance and in theory of propagation.

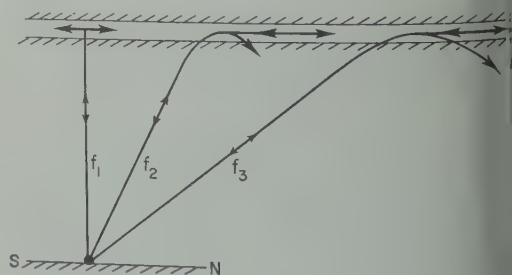
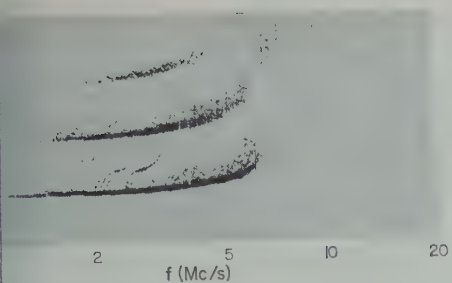


Fig. 13. Ray-path delays increasing with frequency, since waveguide propagation along an elongated irregularity can be excited only near the reflection level.



14. Equatorial striations of unusual shape; the two clearly defined magnetoionic components (the left-hand trace is ordinary, the right-hand trace is extraordinary); Huancayo, 0215 EST, 7, 1960.

ally obvious are certain differences between arctic and equatorial cases, due to the 90° difference in orientation of the irregularities. Among these is the presence of clearly defined striations at comparatively low frequencies in the equatorial case; in the arctic, waveguide penetration and the associated large delays are limited to the higher frequencies. Another marked difference on the equatorial ionograms is the sharp departure of the striations from the main trace. In the arctic case, with the 'holes,' higher frequencies are required for penetration, and the striations are not so

statistical calculation of the conditions for multiple scattering [Bugnolo, 1960] explains the unusual shape of arctic ionograms but does not account for the striations appearing on the example presented. The waveguide model involves multiple scattering, and follows the same general principles, but in addition accounts for the detailed conditions.

EXPERIMENTAL EVIDENCE SUPPORTING THE WAVEGUIDE MODEL

Some experimental observations that support the waveguide model are described in this section. These observations suggest that the equatorial spread F' at the equator is caused by echoes in the vertical north-south direction passing through the ionosonde, whereas the 'equatorial spread F' ' echoes result from scattering in the vertical east-west plane.

In a direction-finding experiment at Huancayo on a fixed frequency of about 8.3 Mc/s, with collinear receiving arrays of dipole antennas and an ionosonde was used with a broad-beam

antenna for transmission). One array had an east-west base line, and the other was north-south, producing fan-shaped beams 7° wide in the north-south or east-west plane, respectively. By suitable phasing of the dipoles, the east-west array could be used to look either east or west at 30° from the vertical, and the north-south array could look 30° north or south. Thus there was a total of six receiver-beam directions, shown schematically in Figure 17. This figure represents a plan view of the ionosphere above the point of observation, and the numbered zones correspond to the six regions of ionosphere examined.

For a pulsed transmission, the amplitude of the returned signal is exhibited as a function of time in the A-scope photographs of Figure 18. (These photographs can be regarded as a vertical section of an ionogram at the 8.3 Mc/s frequency.) The six traces in each sequence correspond to the numbered antenna lobes in Figure 17.

Figure 18a is representative of A-scope recordings taken when there is 'equatorial spread F' ' on the regular ionograms. The almost

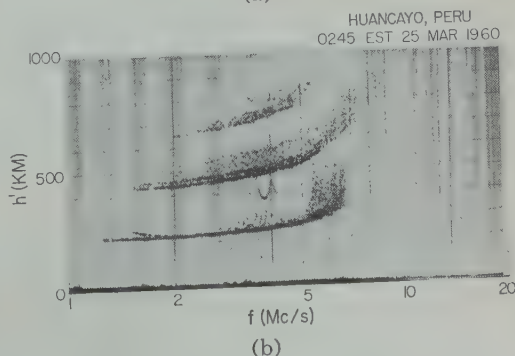
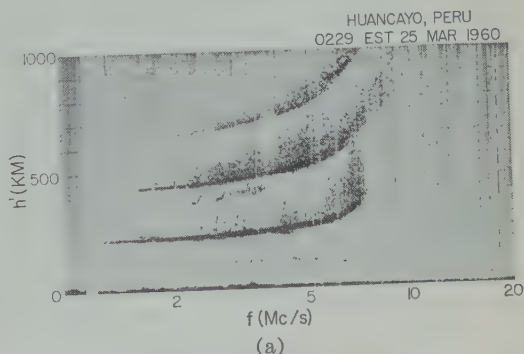


Fig. 15. 'Festoons,' a rare equatorial phenomenon; observed at Huancayo on March 25, 1960: above, at 0229 EST; below, at 0245 EST.

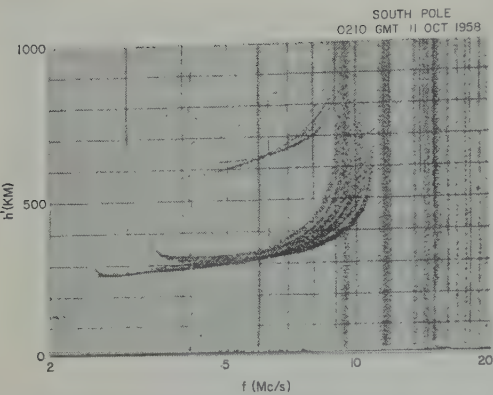


Fig. 16. Spread *F* striations on an antarctic ionogram, which are also explained by waveguide propagation along field-aligned irregularities; south pole, 0210 GMT, October 11, 1958.

complete absence of echoes on traces 3 and 5 clearly demonstrates the east-west nature of the scattered echoes. Figure 18*b*, on the other hand, was obtained during the presence of 'temperate-latitude spread *F*,' and trace 3 exhibits echoes from the south.

In most such observations, the south trace, 3, is stronger than the north trace, 5. Since Huancayo is a little north of the magnetic equator, it is somewhat easier to excite the waveguide modes looking in a southerly direction (Fig. 19). Note also in Figure 18*b* the different structure of the 'temperate-latitude spread *F*' (traces 3 and 5), compared with the 'equatorial spread *F*' (traces 4 and 6).

In some respects, the direction-finding experiments are inconclusive, since the striations are not pronounced at 8.3 Mc/s. At this frequency the return from oblique north-south paths will be weakened by distance, and, as was mentioned in section 4, there may also be some leakage up through the ionosphere. Further, although the 'temperate-latitude spread *F*' is expected mainly from the north and south, its spread in angles of arrival need not be as limited as that of the east-west 'equatorial spread *F*.'

Another feature of Figure 18 is that the time delay of the leading edge of the echoes is only very slightly increased at oblique incidence. It must be remembered that a wave is reflected at lower heights for oblique incidence, so that, for a shallow gradient of electron density, the time delay to the reflection level may not be appreciably increased (Fig. 20).

6. RAY TRACING BELOW THE IRREGULARITIES

It was mentioned in section 4 that the shape of the striations associated with 'temperate-latitude spread *F*' at the equator must be partly due to group delay of the wave packets in the stratified ionosphere below the irregularities. In a recent paper [Calvert and Cohen, 1961] group delay to a fixed height of reflection is calculated as a function of frequency for the ordinary wave and for east-west propagation at the magnetic equator. This work is extended here to take account of the earth's field, with particular emphasis on north-south propagation.

Group propagation at oblique incidence in an anisotropic, plane stratified ionosphere is governed by the four roots, *q*, of the 'Booker quartic' [Booker, 1938]. It is customary to orient the axes so that incidence is confined to the *xz* plane; a plane wave is incident at an angle *θ*, with *S* = sin *θ* and *C* = cos *θ*. The earth's field is represented by a vector **Y** proportional to *bb/f*, with direction cosines (*l, m, n*) corresponding to (*x, y, z*). Then the Booker quartic may be written [Budden, 1961]

$$\begin{aligned} (U - X)\{U(q^2 - C^2) + X\}^2 \\ = Y^2(q^2 - C^2)[U(q^2 - C^2) \\ + X\{1 - (lS + nq)^2\}] \end{aligned} \quad (6)$$

where *U* = 1 - *iv/f*, and *v* is the collision frequency of the ionospheric electrons with the neutral air molecules.

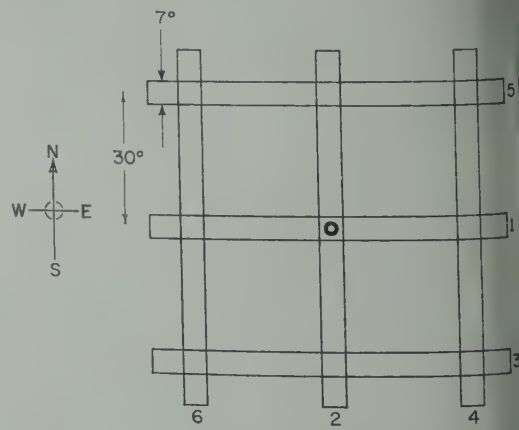


Fig. 17. Schematic representation of the direction-finding antenna used to study spread *F* at Huancayo at about 8 Mc/s; the lobing sequence is indicated by numbers.

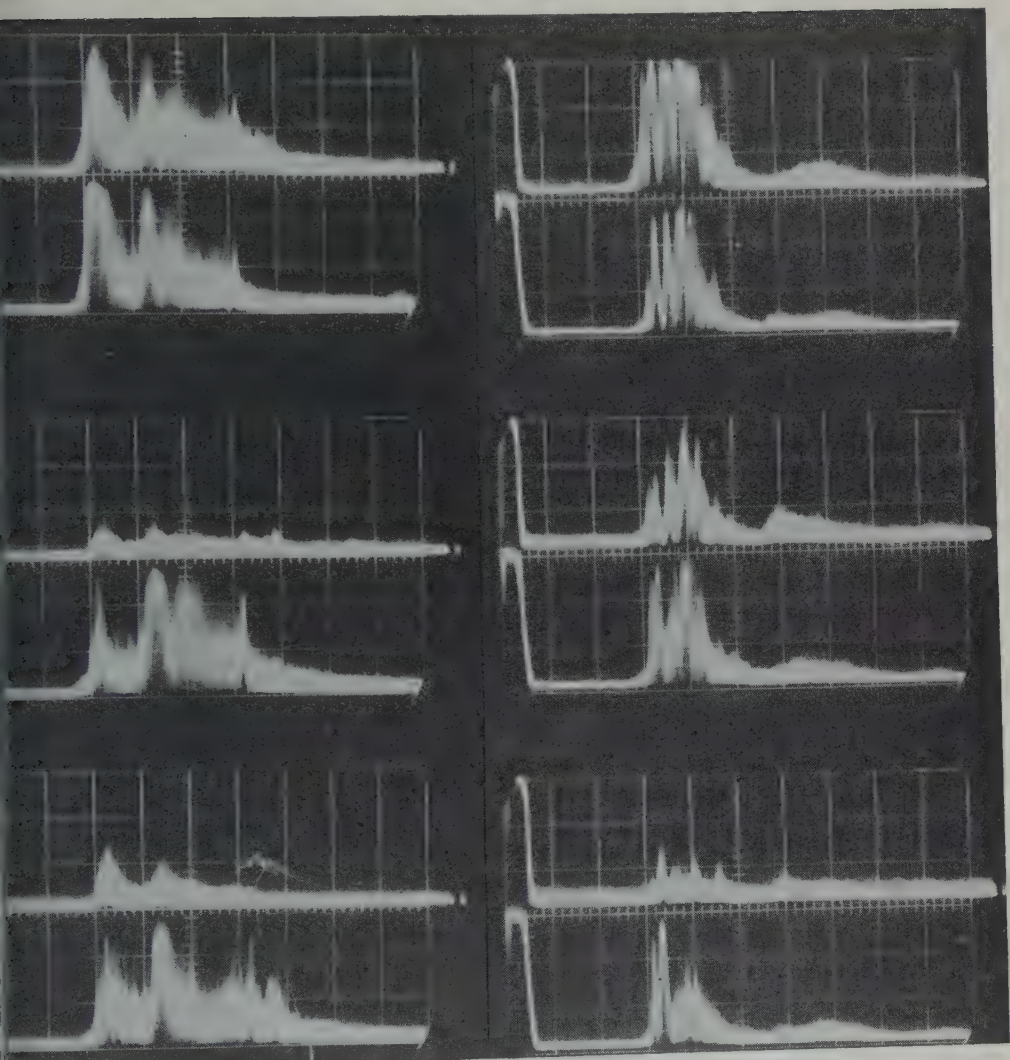


Fig. 18. Direction-finding results at Huancayo; the six traces correspond from top to bottom with the lobing sequence of Figure 17. The amplitude of the echoes appears as a function of their group delay, with a horizontal scale of 150 km per major division. *Left*: East-west echoes at 8.3 Mc/s, indicative of 'equatorial spread F '; 1950 EST, March 15, 1960, north-south polarization. *Right*: Echoes mostly from the south, indicative of 'temperate-latitude spread F ', along with 'equatorial spread F ' from the east and west; 2139 EST, March 16, 1960, east-west polarization.

is simplifies to a biquadratic at the magnetic equator (where the earth's field is horizontal, and $n = 0$). Then reflection occurs when there are two zero roots, so that the term independent of q is equated to zero.

East-west propagation. For east-west propagation at the magnetic equator, $l = 0$, $m = 1$, the Booker quartic becomes

$$(U - X)\{U(q^2 - C^2) + X\}^2 = Y^2(q^2 - C^2)\{U(q^2 - C^2) + X\} \quad (7)$$

The root $q = \pm \sqrt{C^2 - X/U}$ gives the upgoing and downgoing ordinary wave, which is not affected by Y . The other root corresponds to the extraordinary wave.

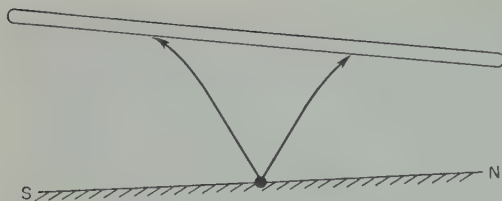


Fig. 19. At Huancayo, slightly north of the magnetic equator, the waveguide modes are more easily illuminated toward the south.

In calculating group delays for reflection at a constant height, $z = h$ (Fig. 13), collisional damping will be ignored. For a given frequency, the angle of incidence must be chosen for reflection at that height. The frequency dependence is given by $X = f_N^2(z)/f^2$, where the 'plasma frequency,' f_N , is proportional to the square root of the electron number density, N , and $Y = f_H/f$, where f_H is the 'gyrofrequency,' about 0.73 Mc/s at 250 km height in the equatorial F layer.

A root, q , of the Booker quartic describes the phase propagation of a wave packet inside the ionosphere with spatial and temporal variation given by

$$\left\{ -2\pi i f \left(\frac{Sx}{c} + \frac{1}{c} \int q \, dz - t \right) \right\}$$

where c is the velocity of light in vacuo. Then from this dependence the group velocity, ray path, and thus 'virtual height' (corresponding to one-half the time delay) can be determined. For the east-west ordinary polarization, reflection at $z = h$ requires $C^2 = f_N^2(h)/f^2$, and the virtual height, h' , is given by

$$h' = f \int_0^h \frac{dz}{\sqrt{f_N^2(h) - f_N^2(z)}} \quad (8)$$

Thus h' is proportional to f for any electron-density profile $f_N(z)$.

For the east-west extraordinary polarization, the condition for reflection is that

$$C^2 = \frac{X(1-X)}{1-X-Y^2} \quad (9)$$

To evaluate h' for this case is more difficult, because of the presence of Y . However, an approximation is available for obtaining $h'(f)$, applicable to the north-south case as well. If most of the time delay is imposed in free space

below the ionosphere (Fig. 21), h' is given by

$$h' \approx h/C \quad (10)$$

If this approximation is used for the east-west ordinary case, (8) is replaced by

$$h' \approx \frac{fh}{f_N(h)} \quad (11)$$

This relation has the correct frequency dependence, lending some confidence to the approximation.

At $z = h$, the reflection condition (9) defines C as a function of frequency for the extraordinary ray, leading to

$$h' \approx \frac{h}{C} = \frac{fh}{f_N} \sqrt{1 - \frac{f_H^2}{f^2 - f_N^2}} \quad (12)$$

where f_N is again evaluated at $z = h$.

An example for both ordinary and extraordinary rays is exhibited in Figure 22, where h'/C is plotted as a function of $\log f$.

North-south propagation. For north-south propagation at the magnetic equator, $l = 1$, and the Booker quartic becomes

$$(U - X)\{U(q^2 - C^2) + X\}^2 = Y^2(q^2 - C^2)\{U(q^2 - C^2) + XC^2\} \quad (13)$$

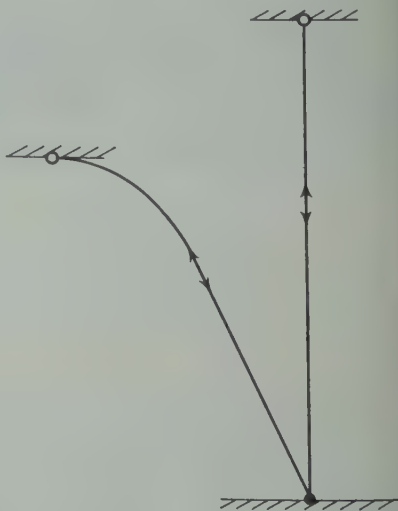
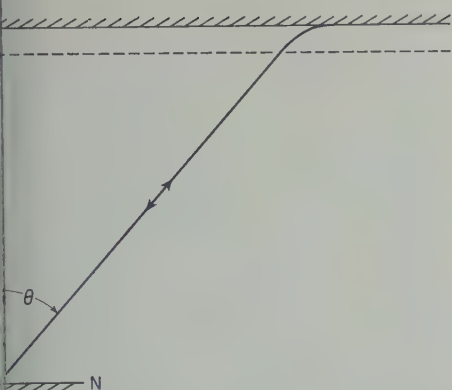


Fig. 20. With a shallow gradient of electron density, the oblique incidence scatter from reflection level irregularities occurs lower down for a given frequency and therefore need not be delayed much more than the normal-incidence reflection.



21. If most of the ray path lies in free space, group delay to the reflection level at oblique incidence is given by $h' \approx h/C$, where $C = \cos \theta$.

the additional C^4 in the last term of (on 7). Without damping, reflection occurs at $X = 1$ or $X = C^2(1 \pm Y)$ (Fig. 23). For C^2 up to $C^2 = 1/(1 + Y)$, marked by an arrow, the ordinary wave is reflected at $X = 1$. There is a transition, associated with the change from quasi-transverse to quasi-longitudinal propagation [Ratcliffe, 1959] (the curves are drawn as though resolved by the presence of collisional damping). Also indicated is the 'plasma resonance' level at $X = 1 - Y^2$,

where induced space charges would give rise to large electric wave fields for the extraordinary component, had it not already been reflected lower in the ionosphere. In addition, there is critical coupling between the ordinary and extraordinary waves at the level where $X = 1 + Y^2 C^4 / 4S^2$ [Pitteway, 1959], but again this happens too high in the ionosphere to affect the present problem.

For $X = 1$, $f = f_N(h)$ for all values of C and h' . For reflection at $X = C^2(1 \pm Y)$,

$$\frac{h}{C} = \frac{hf}{f_N(h)} \sqrt{1 \pm \frac{f_H}{f}} \quad (14)$$

Curves corresponding to Figure 22 are plotted for north-south propagation in Figure 24. Note the constant frequency separation, f_H , of the two curves beyond the $C^2 = 1/(1 + Y)$ transition.

Comparison of ray tracing with ionograms. The striations on the ionograms of Figure 5 have about twice the slope of those predicted by Figure 24, and so they cannot be accounted for solely by ray tracing to horizontal patches of irregularities. However, the magnetic field lines at Huancayo are curved slightly downward relative to the earth's surface, with radius of curvature about one-half that of the earth. Thus, field-aligned irregularities, which at the

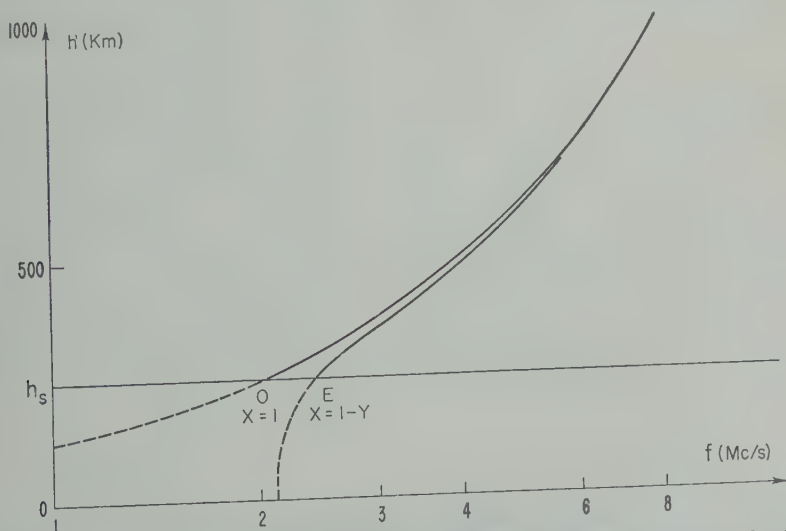


Fig. 22. Approximate group delay for oblique incidence to a constant height of reflection, for east-west propagation at the magnetic equator. The irregularities are situated at a height $h_s = 250$ km, where the plasma frequency is assumed to be 2 Mc/s and the gyrofrequency is 0.73 Mc/s. The ordinary delay is proportional to frequency, and the extraordinary delay approaches it for higher frequencies.

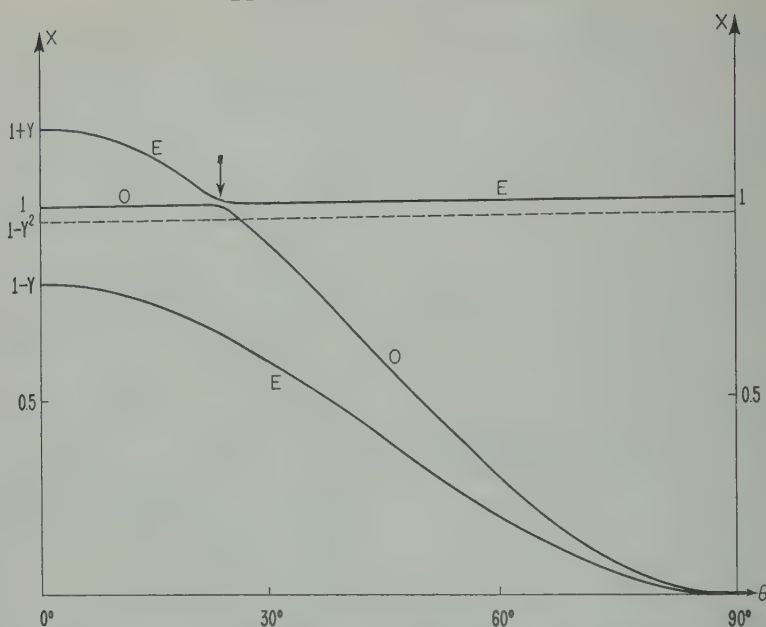


Fig. 23. The reflection levels as a function of angle, for oblique incidence in the north-south plane at the magnetic equator. Note the transition (marked by an arrow) of the ordinary and upper extraordinary reflection levels. The curves correspond to $Y = 0.2$ (a wave frequency of 3.65 Mc/s); the broken line represents the plasma resonance level, where $X = 1 - Y^2$.

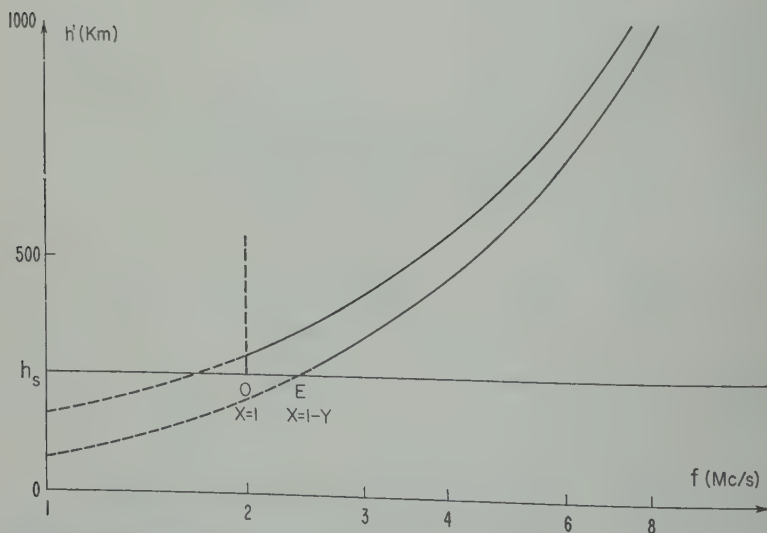


Fig. 24. Curves corresponding to Figure 22, this time for north-south propagation at the magnetic equator. The vertical segment of the ordinary trace results from the 'Spitze' phenomenon (constant reflection at $X = 1$ for angles up to $C^2 = 1/(1 + Y)$), but is not observed on the ionograms. The curved segment of the ordinary trace lies at a constant spacing, f_H , from the extraordinary trace.

equator are at a height of 250 km, at 25 km lower at a distance of 400 km from the equator, and a lower frequency will be observed with reflection at this lower height. This geometry will steepen the curves of Figure 10, especially when the main trace on ionograms implies a steep gradient of electron density. Furthermore, the effect of retardation inside the waveguide may also be observed in the traces.

One apparent discrepancy is that the vertical separation of the two magnetoionic traces in Figures 5 and 14 is only about half the ordinary frequency, and the ordinary trace has no vertical portion. The vertical portion of the trace in Figure 24 is due to a reflection at $C^2 = 1$ that is constant over a range of angles from the vertical, and this is associated with the 'Spitze' phenomenon [Poevlein, 1950]. As mentioned in section 4, the ordinary ray does not become perpendicular to the earth's magnetic field, and, for angles less than $C^2 = 1/C^2$, this tendency produces a cusp at the turning point in the ordinary ray path. Perhaps retardation inhibits excitation of the waveguide irregularities in the 'Spitze' region, since, although the wave-normal is pertinent to the propagation in the waveguide, the cusp may prevent energy from entering the mode. Therefore, it is not surprising that none of the observed traces exhibit a vertical segment.

Before and after a wave is reflected in the F region it must penetrate the E region and the F region. Thus the $h' \approx h/C$ assumption in section 4 is very approximate. To improve this approximate result would require a model of actual ionosphere profiles, in which the F layer is at least as thick as the free space below it. In this case, computer integrations would be required to give $h'(f)$. It is to be hoped that for a model of the ordinary and extraordinary ray separation might be reduced from f_H to $f_H/2$, without the appearance of the 'Spitze.'

7. CONCLUSIONS

The approximate theory of this paper accounts for the configurations on equatorial ionograms observed as 'temperate-latitude spread F ,' and also explains the presence of striations in these configurations. This class of spread F is caused by echoes from irregularities with large

scale sizes, which are effective only at the reflection height in the ionosphere, and which scatter in the north-south plane by a waveguide process.

This is the third spread- F category which can be attributed to various kinds of irregularities in the equatorial ionosphere: (i) 'satellites' resulting from large scale deformations of the isoionic contours; (ii) 'equatorial spread F ' resulting from thin, elongated irregularities; (iii) 'temperate-latitude spread F ' resulting from much thicker, elongated irregularities. Class ii irregularities support oblique scatter at 50 Mc/s, whereas those of class iii do not. Class ii echoes arrive in the east-west plane, whereas class iii echoes arrive in the north-south plane passing through the ionosonde. Class ii irregularities scatter effectively at all sounder frequencies, whereas class iii irregularities are only effective at the reflection level, and give a striated structure similar in appearance and interpretation to that of 'arctic spread F .'

The waveguide model for the formation of striations requires irregularities with a vertical extent greater than about 250 meters. The frequency range of the individual striations indicates that the irregularities have horizontal correlation over distances as great as 200 km along the magnetic field lines. The east-west variation of the irregularities has not been treated explicitly.

Acknowledgments. The experimental collaboration of K. L. Bowles and G. R. Ochs of the National Bureau of Standards and of A. A. Giesecke, Jr., Director-Técnico, M. Tábara, J. Lanat, and H. Goller of the Instituto Geofísico del Perú is gratefully acknowledged. Discussions with W. Calvert, T. N. Gautier, J. T. Brown, Jr., and J. W. Wright of the National Bureau of Standards were of considerable help. M. L. V. Pitteway is indebted to Harkness House for a maintenance grant while he was a guest worker at the National Bureau of Standards.

REFERENCES

- Booker, H. G., Propagation of wave-packets incident obliquely upon a stratified doubly refracting ionosphere, *Phil. Trans. Roy. Soc. London, A*, **237**, 411-451, 1938.
- Booker, H. G., and M. L. V. Pitteway, Waveguide propagation down elongated irregularities in the ionosphere, in press, 1961.
- Bowles, Kenneth L., and Robert Cohen, A study of

- radio wave scattering from sporadic *E* near the magnetic equator, *Ionospheric Sporadic E*, Pergamon Press, New York, in press, 1961.
- Budden, K. G., *Radio Waves in the Ionosphere*, Cambridge University Press, 542 pp., 1961.
- Bugnolo, Dimitri S., Spread *F* and multiple scattering in the ionosphere, *J. Geophys. Research*, **65**, 3925-3929, 1960.
- Calvert, Wynne, and Robert Cohen, The interpretation and synthesis of certain spread-*F* configurations appearing on equatorial ionograms, *J. Geophys. Research*, **66**, 3125-3140, 1961.
- Cohen, Robert, and Kenneth L. Bowles, On the nature of equatorial spread *F*, *J. Geophys. Research*, **66**, 1081-1106, 1961.
- Pitteway, M. L. V., The reflexion of radio waves from a stratified ionosphere modified by weak irregularities, *Proc. Roy. Soc. London, A*, **252**, 556-569, 1958.
- Pitteway, M. L. V., Reflexion levels and coupling regions in a horizontally stratified ionosphere, *Phil. Trans. Roy. Soc. London, A*, **252**, 53-103, 1959.
- Pitteway, M. L. V., The reflexion of radio waves from a stratified ionosphere modified by weak irregularities, *2, Proc. Roy. Soc. London, A*, **252**, 86-100, 1960.
- Poevverlein, H., Strahlwege von Radiowellen in der Ionosphäre, *Z. angew. Phys.*, **2**, 152-160, 1959.
- Ratcliffe, J. A., *The Magneto-Ionic Theory and Applications to the Ionosphere*, Cambridge University Press, 206 pp., 1959.

(Manuscript received July 17, 1961.)

On the Index of Refraction in the Ionosphere

OTTO THEIMER AND LEONARD S. TAYLOR¹

New Mexico State University, University Park, New Mexico

Abstract. The forces acting upon an electron in an ionized gas are investigated to determine the index of refraction of the medium. The result previously established by Darwin that the force due to the polarization field in an ionized gas of low concentration is canceled by the effect of collisions between ions and electrons is re-established in a simple manner that clearly displays the physical basis for the effect. Because of this cancellation the Sellmeyer formula,

$$n^2 - 1 = -Ne^2/\epsilon_0 m \omega^2$$

determines the relation between the electron number density, N , and the index of refraction, n , in the ionosphere when collisions are neglected. It is shown that this result is independent of the effects of free-electron orbits which penetrate the orbits of bound electrons and of the effect of asymmetric charge distributions in the ions. The breakdown of the theory at high frequencies is indicated.

Introduction. The question of the correct formula for the index of refraction of an ionized gas has had a curious role in the history of physics. Although the principal arguments were made in the nineteenth century, it was not until 1943 that a theory was established by Darwin [1943] which determined that the correct formula for the index of refraction of the ionosphere must be based upon the Sellmeyer

$$n^2 - 1 = -Ne^2/\epsilon_0 m \omega^2 \quad (1)$$

or the Lorentz formula,

$$n^2 - 1/(n^2 + 2) = -Ne^2/\epsilon_0 m \omega^2 \quad (2)$$

The distinction between these two formulas does not depend on any of the theories of collision which result in the presence of a collision frequency in the formulas determining the index of refraction at low frequencies. The distinction is made on the basis of whether the polarization force must be accounted for in an ionized gas. The question of this type seemed to be answerable on the basis of first principles, and no appeal to experiment appeared necessary. Because of the controversy about the manner in which the Lorentz formula should be applied in this case, however, an appeal was made to experiment. Unfortunately, experiment yielded inconclusive results [Mitra, 1952]. It was against this

background that Darwin finally succeeded in proving that the Sellmeyer formula, equation 1, is the correct expression by demonstrating that the radiation-induced perturbations of the orbits of the free electrons produce an effect equivalent to that produced by a force that exactly cancels against the force due to the polarization field. Although no specific objections to Darwin's analysis have been raised, the result has been accepted only with considerable caution in the literature [Mitra, 1952; Ratcliffe, 1959], probably because of the highly mathematical nature of the solution, which leaves unanswered certain questions about the physical basis of the depolarizing force. A number of other puzzling features also appear in the Darwin paper which need not be discussed here. In the present work the authors hope to clarify this situation with an independent demonstration of the validity of the Darwin result which enjoys the advantage of extreme simplicity and physical clarity. In addition, certain results not previously proved are obtained here. Finally, the disappearance of the depolarizing force at high frequencies is predicted, and the frequencies at which the Lorentz formula becomes valid are estimated.

Forces in an ionized gas. We begin the analysis by adopting the methods used in describing the forces acting on a molecule in an ordinary dielectric [Panofsky and Phillips, 1955]. In order to employ these methods, it is only necessary to consider that the motion of the electrons is compounded of the motion of a 'guiding center' plus a radiation-induced small

¹ Present address: Missile and Space Vehicle Department, General Electric Company, Philadelphia, Pennsylvania.

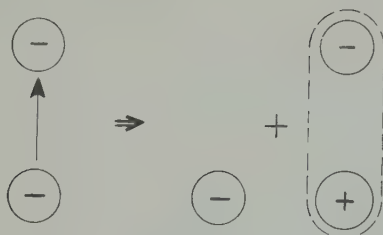


Fig. 1. Formation of a virtual dipole.

displacement from the guiding center. This small displacement is equivalent to the addition of a virtual dipole as illustrated in Figure 1. The 180° phase shift of the free-electron oscillation with respect to the phase of the applied field results in a polarization field directed against the applied field, but this change in sign is immaterial to the theory. The ions may be assumed to have zero mobility in the external field because of their relatively large mass, and they are fixed in the reference frame. In mks units, the forces acting on a free electron in the gas may be described as follows:

$\mathbf{F}_1 = -e\mathbf{D}/\epsilon_0$ is the force produced by the sources of the external field which may be regarded as the true charges on the surface of a parallel-plate capacitor. (The electron charge is taken to be $-e$.)

$\mathbf{F}_2 = e\mathbf{P}/\epsilon_0$ is the force produced by the virtual polarization charge on the surface of the medium facing the capacitor plates.

$\mathbf{F}_3 = -e\mathbf{P}/3\epsilon_0$ is the force due to the virtual polarization charge induced on the surface of an imaginary auxiliary sphere centered at the position of the test electron. The radius of the auxiliary sphere is arbitrary, but to avoid unnecessary complications, it is specified that the radius must be small compared with the wavelength of the radiation field.

\mathbf{F}_4 is the force due to the virtual dipoles within the auxiliary sphere. Because these are randomly distributed, the ensemble average of this force is zero.

\mathbf{F}_5 is the force due to the ions and the undispersed electrons. It is known [Theimer and Gentry, 1959] that an excellent approximation to this force may be obtained on the assumption that the test electron is subject to the Coulomb field of the nearest neighbor ion. (The other charges contribute forces which affect the field most strongly at distances from the nearest

neighbor ion which approach half the mean interionic spacing. It will be seen that \mathbf{F}_5 is strongly dependent on the field in the neighborhood of the nearest neighbor ion, so that the use of the Coulomb field in this connection is justified. It may also be remarked here that the Debye radius is large compared with the mean interionic spacing for the conditions encountered in the ionosphere, so that it is not necessary to account for any shielding of the Coulomb field by statistical correlations of the other charges. The average of \mathbf{F}_5 is zero in the absence of an external field, because of the isotropic distribution of the orbits of the free electrons. In the presence of an external radiation field, it will be shown that $\mathbf{F}_5 = -\mathbf{F}_3$.

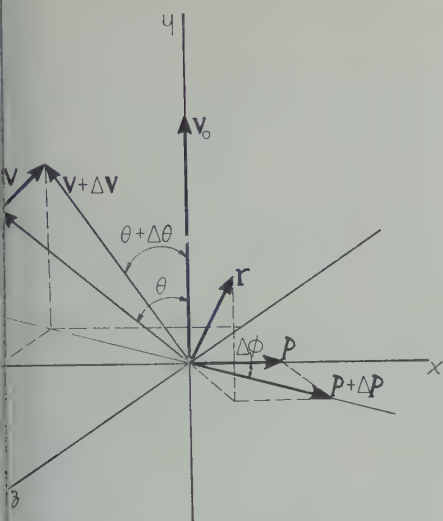
\mathbf{F}_6 is the force produced by induced atomic and ionic dipoles. This force may be obtained in the conventional manner. It suffices for the purposes of this paper to assume that the neutral atoms and molecules and the ions have zero polarizability, so that $\mathbf{F}_6 = 0$.

The depolarizing force in an ionized gas. The effect of the radiation-induced displacement upon the orbit of a free electron during a single encounter with an ion will now be calculated. The radiation-induced perturbation of the electron velocity will be neglected for the moment; it will be shown later to have negligible effect. The frequency of the external field is assumed to be such that

$$\omega/2\pi \ll v/d \quad (1)$$

where d is the mean interionic spacing and v is the electron thermal velocity. Thus the displacement effected by the field has occurred over a period long compared with the time of passage through the single nearest neighbor zone under examination. During the latter period the effect of the external field is neglected. It is only assumed that the electron enters the zone with some infinitesimal displacement, \mathbf{r} . The ensemble average of this quantity is reasonably assumed to be equal to the average displacement of the other electrons in the gas which produces the virtual dipoles that create \mathbf{F}_3 .

Figure 2 displays the geometry of the interaction situation. An electron, velocity \mathbf{v}_0 , directed along the local y axis, collision parameter \mathbf{b} directed along the local x axis, encounters an ion, charge Ze , located at the origin. In the absence of an external field, the electron would



Geometry of the collision perturbation.

the collision zone with velocity \mathbf{v} and on θ . As a result of the perturbation, \mathbf{r} , corresponding change of \mathbf{p} the electron with velocity $\mathbf{v} + \Delta\mathbf{v}$ and deflection $\theta + \Delta\theta$.

apparent that the perturbation of the position along \mathbf{v}_0 is of no consequence; perturbation of the magnitude of \mathbf{p} changes magnitude of the deflection angle; the rotation perpendicular to \mathbf{p} and \mathbf{v}_0 rotates scattering plane through an angle $\Delta\phi$. Thus,

$$\mathbf{p} + \Delta\mathbf{p} \approx (p + r_x)\hat{x} \quad (4)$$

the caret denotes a unit vector, and

$$\mathbf{v} = -v(\sin \theta)\hat{x} + v(\cos \theta)\hat{y} \quad (5)$$

is that

$$\begin{aligned} \mathbf{v} &= -v \sin(\theta + \Delta\theta) \cos(\Delta\phi)\hat{x} \\ &\quad + v \cos(\theta + \Delta\theta)\hat{y} \\ &\quad - v \sin(\theta + \Delta\theta) \sin(\Delta\phi)\hat{z} \\ &\approx -v[\sin \theta - \Delta\theta \cos \theta]\hat{x} \\ &\quad + v[\cos \theta - \Delta\theta \sin \theta]\hat{y} \\ &\quad - v \Delta\phi (\sin \theta)\hat{z} \end{aligned} \quad (6)$$

using

$$\Delta\phi \approx \tan(\Delta\phi) = r_z/p \quad (7)$$

change of velocity is

$$\begin{aligned} \Delta\mathbf{v} &= -v \Delta\theta (\cos \theta)\hat{x} \\ &\quad - v \Delta\theta (\sin \theta)\hat{y} - v r_z (\sin \theta)\hat{z}/p \end{aligned} \quad (8)$$

But to the first order

$$\Delta p = r_z \quad (9)$$

Thus, with some manipulation,

$$\begin{aligned} \Delta\mathbf{v} &= -v \frac{d(\sin \theta)}{dp} r_x \hat{x} \\ &\quad + v \frac{d(\cos \theta)}{dp} r_x \hat{y} - v \frac{\sin \theta}{p} r_z \hat{z} \end{aligned} \quad (10)$$

Since the average velocity change must be parallel to \mathbf{E} , the desired quantity is the component of $\Delta\mathbf{v}$ in the \mathbf{r} direction. Thus, taking the scalar product of $\Delta\mathbf{v}$ with \mathbf{r} and dividing by r ,

$$\begin{aligned} (\Delta\mathbf{v})_r &= -v \frac{d(\sin \theta)}{dp} \left(\frac{r_x^2}{r} \right) \\ &\quad + v \frac{d(\cos \theta)}{dp} \left(\frac{r_x r_y}{r} \right) - v (\sin \theta) r_z^2 / pr \end{aligned} \quad (11)$$

Next, it is necessary to average $(\Delta\mathbf{v})_r$ over all relative directions of the electron velocity and impact parameter corresponding to a given direction of \mathbf{r} in space. Some care must be exercised in this process, because, for any given direction of \mathbf{p} , \mathbf{v} may assume any direction in the plane perpendicular to \mathbf{p} . The geometry of the situation is illustrated in Figure 3. A triad (X, Y, Z) is chosen which is fixed in space with the Z axis parallel to \mathbf{r} . Oz' is the intersection of the yz plane with the XZ plane, and Oy' is the intersection of the yz plane with XY plane. Thus

$$(r_x^2)_{AV} = (r^2 \cos^2 A)_{AV} = r^2/3 \quad (12)$$

Next, using

$$\cos B = \cos C \cdot \sin A \quad (13)$$

it follows that

$$(r_z^2)_{AV} = r^2 (\cos^2 C)_{AV} \cdot (\sin^2 A)_{AV} = r^2/3 \quad (14)$$

Obviously

$$(r_x r_y)_{AV} = 0 \quad (15)$$

Using equations 12, 14, and 15, it is possible to combine the terms in equation 10, with the result

$$[(\Delta\mathbf{v})_r]_{AV} = - \left(\frac{vr}{3p} \right) \frac{d(p \sin \theta)}{dp} \quad (16)$$

Next, it is necessary to average equation 16 over all values of the collision parameter from zero to some maximum value p_m . Because of the form in which equation 16 has been cast, this average is immediately obtainable. Thus

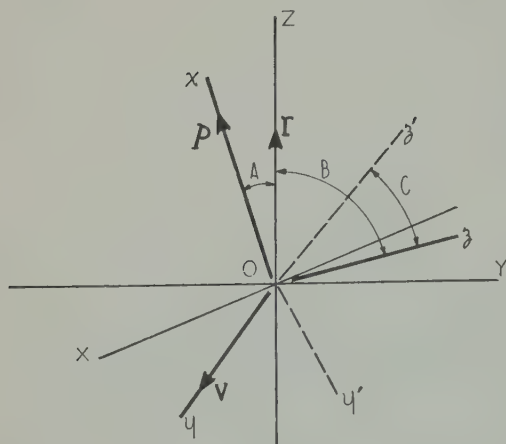


Fig. 3. Geometry of the averaging process.

$$\begin{aligned} \overline{(\Delta \mathbf{v})}_r &= \frac{2}{p_m^2} \int_0^{p_m} p [(\Delta \mathbf{v})_r]_{AV} dp \\ &= -\left(\frac{2vr}{3p_m}\right) \sin \theta_m \quad (17) \end{aligned}$$

where θ_m is the deflection suffered by an electron with collision parameter p_m .

During a time interval Δt , the number of encounters between an electron and an ion with a collision parameter p_m or smaller is $\pi N_i p_m^2 v \Delta t$, where N_i is the ion concentration ($N_i = N/Z$). Therefore, the average velocity change during Δt is

$$\overline{(\Delta \mathbf{v})}_r = -(2\pi N_i v^2 \Delta t / 3) \cdot p_m \sin \theta_m \quad (18)$$

This velocity change may be described as the 'result' of the depolarizing force, \mathbf{F}_5 , where

$$\mathbf{F}_5 = -(2\pi N_i v^2 m / 3) \cdot p_m \sin \theta_m \quad (19)$$

and m is the electron mass. In the next section a description of the geometry of the interaction orbits will be given which will clarify the meaning of equation 19. In the meantime, it is useful to note that, if it were assumed that all encounters were such that the electrons experienced the complete Rutherford scattering deflection, then, since in this case [Goldstein, 1953]

$$p_m \sin \theta_m = 2p_0 / (1 + p_0^2 / p_m^2) \quad (20)$$

it follows that

$$\begin{aligned} \mathbf{F}_5 &= -4\pi N_i p_0 v^2 m / 3 \\ &= -4\pi N p_0 v^2 m / 3Z \quad (21) \end{aligned}$$

Here use had been made of the fact that the collision parameter corresponding to a 90°

deflection,

$$p_0 = Ze^2 / 4\pi\epsilon_0 mv^2 \quad (22)$$

is some thousands times less than the mean interionic spacing for all conditions of present interest, so that the condition

$$p_0 \ll p_m \quad (23)$$

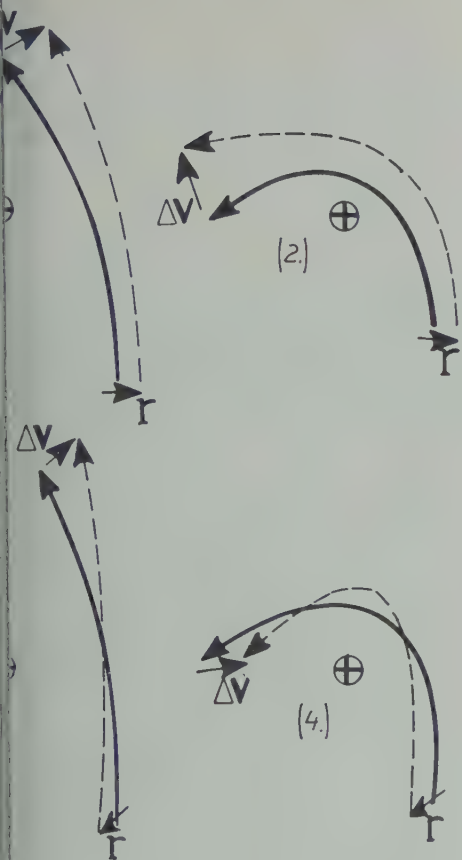
may be applied to equation 20. Combining equations 21 and 22,

$$\mathbf{F}_5 = -Ne^2 \mathbf{r} / 3\epsilon_0 = e\mathbf{P} / 3\epsilon_0 = -\mathbf{F}_3 \quad (24)$$

Description of perturbed orbits. To employ the result expressed in equation 19 correctly, it is necessary to consider the geometric basis of this formula, which is independent of any particular scattering law to be employed. It is apparent from equation 11 that the depolarizing force is capable of analysis into a force due to change of deflection angle plus one due to rotation of the plane of deflection. These are represented by the first and last terms in equation 11, respectively. (The second term averages to zero.) In Figure 4 some representative orbits are shown which display the sources of the forces. It is apparent from (1) and (2) of Figure 4 that the force resulting from the change of deflection angle changes sign at the collision parameter corresponding to a 90° deflection, and that $\Delta \mathbf{v} \cdot \mathbf{r}$ is positive for deflections greater than 90° . On the other hand, it may be seen from (3) and (4) of Figure 4 that the force resulting from the rotation of the plane of the deflection does not change sign, and that $\Delta \mathbf{v} \cdot \mathbf{r}$ is negative. Moreover, it is readily shown that, because of the long range of Coulomb interactions, the average of each of the component forces diverges logarithmically when equation 11 is evaluated for an inverse-square law. Because of the cancellation of the two component forces for large collision parameters, however, the total depolarizing force tends to zero much more rapidly than a simple Coulomb force. In fact, equation 19 could have been written

$$\mathbf{F}_5 = -\left(\frac{2\pi N_i v^2 m}{3}\right) \int_0^{p_m} \frac{4(p/p_0) dp}{(1 + p^2/p_0^2)^2} \quad (25)$$

where the Rutherford scattering law has been employed and the integration over p not explicitly carried out. In this form, which is exactly the same as that achieved by Darwin after the variable of integration has been changed from the hyperbolic eccentricity to the collision parameter,



4. Collision trajectories: (1) \mathbf{r} parallel to \mathbf{p} , $p < p_0$; (2) \mathbf{r} parallel to \mathbf{p} , $p < p_0$; (3) \mathbf{r} perpendicular to \mathbf{p} , $p > p_0$; (4) \mathbf{r} perpendicular to \mathbf{p} , $p < p_0$.

It is apparent that the dominant contribution to the depolarizing force is due to electrons with small collision parameters. Approximately 1 per cent of the total force is contributed as a result of encounters with collision parameters less than $5 p_0$, corresponding to deflections of 23° or greater. Because p_0 is so much smaller than the mean interionic spacing, electrons involved in these close encounters may legitimately be assumed to experience the complete Rutherford deflection. Thus the use of the complete Rutherford scattering equation 19 is correct, and equation 19 shows that the polarization field is canceled by the effect of collisions in the ionosphere. It has not been necessary, of course, to consider the effect of collisions with neutral atoms in this section, because the mean free path of the electrons through the neutral gas is many times greater than the mean interionic spacing in all

cases of practical interest. Note that it has not been necessary to consider the fact that electrons with collision parameters approaching one-half the mean interionic spacing will not experience the complete deflection. Thus the present theory avoids the difficulties inherent in any attempt to define a 'shape' for the interaction zone and in the computation of the scattering law for an electron passing through such an interaction zone. This artificiality exists in the Darwin presentation, where it is shown that a reasonable choice for the interaction zone is a sphere of radius b , such that

$$p_0 \ll b \ll d \quad (26)$$

It is the computation of the scattering law for this restricted zone of interaction that is the source of the mathematical complexity of the theory as originally presented.

It has not yet been shown that the effects of perturbations of the electron velocities can be ignored in the analysis. This may be seen in the following manner:

$$|\Delta\theta|_v = \frac{4(p/p_0)}{1 + p^2/p_0^2} \frac{|\Delta v|}{v} \quad (27)$$

is the magnitude of the increment in scattering angle due to an increment of velocity, $|\Delta v|$, using the Rutherford scattering law. The increment due to a change of collision parameter, $|\Delta p|$, is

$$|\Delta\theta|_p = \frac{2/p_0}{1 + p^2/p_0^2} |\Delta p| \quad (28)$$

Using

$$|\Delta v| = \omega |\Delta p| \quad (29)$$

the ratio

$$|\Delta\theta|_v/|\Delta\theta|_p = 2p\omega/v \ll 1 \quad (30)$$

where use has been made of the assumption

$$p_m \ll v/2\omega \quad (31)$$

This condition is satisfied for all frequencies of interest in ionospheric radio research. Thus the force arising from the perturbation of the electron velocity may be neglected. We note in passing that this force is of the viscous type and can always be handled quite separately [Theimer and Taylor, 1960] from the question of the cancellation of the polarization field.

Conclusions. The material in the preceding sections is equivalent to the proof given by Darwin that the polarization field need not be accounted for in an ionized gas. As a result, the

equation of motion of an electron is given by

$$m\ddot{\mathbf{r}} = \mathbf{F}_1 + \mathbf{F}_2 = -e\mathbf{E}_0 \exp(i\omega t) \quad (32)$$

It then follows that the current is

$$\mathbf{J} = (-iNe^2/m\omega)\mathbf{E}_0 \exp(i\omega t) \quad (33)$$

so that the conductivity is given by

$$\sigma = -Ne^2/m\omega \quad (34)$$

Thus, in a region of unit specific magnetic inductive capacity,

$$n^2 - 1 = -i\sigma/\epsilon_0\omega = -Ne^2/\epsilon_0m\omega^2 \quad (35)$$

the Sellmeyer formula.

The generality of the result expressed in equation 19 permits certain other conclusions to be drawn which strengthen the conviction that the Sellmeyer expression is the correct formulation for the index of refraction in the ionosphere:

First, the result is independent of the charge of the ions. (We always assume a neutral gas, of course.) This result also appeared in the original work of Darwin. It is an obvious consequence of equations 21 and 22.

Second, it may be seen from equation 19 that the result is independent of variations from the Rutherford scattering law at small collision parameters. Thus the effect of free-electron orbits which penetrate the orbits of bound electrons may be neglected. In this connection, we naturally do not consider recombination processes, or any other processes, that would make the collision inelastic.

Third, because the interionic spacing is so much larger than any atomic or molecular dimension in the circumstances under consideration, it is always possible to choose p_m sufficiently large that the effects of dipole and higher-order moments upon the scattering process may be neglected in comparison with the monopole deflection. Thus the effects of asymmetric distributions of charge, such as exist in ionized molecules, need not be considered.

Fourth, it is apparent that the condition expressed in equation 3 is unnecessarily strong. It is only necessary to assure that the displacement, \mathbf{r} , has been effected by the radiation field over a period long compared with that during which electrons with collision parameters equal to $5p_0$ or less can effect essentially the complete Rutherford deflection. An excellent approximation to the orbit of an electron with a collision parameter of $5p_0$ may be obtained by assuming

that the electron moves subject to the action of the time-dependent force it would experience if it maintained its original straight-line path. It then follows quite simply that the approximate deflection is given by integrating this force from $-t_m$ to t_m , with the result

$$\sin \theta_{5p_0} \approx \frac{2}{5} \frac{vt_m}{[(5p_0)^2 + (vt_m)^2]^{1/2}} \quad (36)$$

Thus it may be required that the period of the radiation field only be larger than $v/50$. This requirement is fulfilled in the radio portion of the spectrum, but in the infrared and visible regions it is expected that the Sellmeyer formula will no longer hold true and that the Lorentz formula will become valid. This results as the frequency of the field becomes high enough that the electrons tend to follow the guiding center of their undisturbed paths through the ionic force fields and the depolarizing force vanishes. The authors intend to investigate this transition region in detail in a later work.

Finally, the strange cancellation of the polarization field by the effect of collisions may be considered to be another of the remarkable features of the inverse square law. This feature is no more remarkable, however, than the fact that it is only possible to define a polarization field in the manner described here for an inverse square law of force because it is only in this case that the radius of the auxiliary sphere upon which the polarization charges are imagined to appear is arbitrary.

REFERENCES

- Darwin, C. G., The refractive index of an ionized medium, II, *Proc. Roy. Soc. London, A*, 18, 152-166, 1943.
- Goldstein, H. *Classical Mechanics*, Addison-Wesley Publishing Co., Cambridge, 1953.
- Mitra, S. K., *The Upper Atmosphere*, Asiatic Society of Calcutta, 1952.
- Panofsky, W. K. H., and M. Phillips, *Classical Electricity and Magnetism*, Addison-Wesley Publishing Co., Cambridge, 1955.
- Ratcliffe, J. A., *The Magneto-Ionic Theory and Applications to the Ionosphere*, Cambridge University Press, Cambridge, 1959.
- Theimer, O., and R. Gentry, Scattering potentials in fully ionized gases, *Phys. Rev.*, 116, 787-79, 1959.
- Theimer, O., and L. S. Taylor, The frequency dependence of the coefficient of dynamic friction of a plasma, *Ann. Phys.*, 11, 377-391, 1960.

(Manuscript received March 24, 1961;
revised July 15, 1961.)

Solar-Flare Effects on 2.5 and 5.0 Mc/s Atmospheric Radio Noise

JOHN R. HERMAN

*Research and Advanced Development Division, AVCO Corporation
Wilmington, Massachusetts*

Abstract. Analysis of radio noise records from Kekaha, Hawaii, and Ohira, Japan, during 75 solar flares occurring during August, September, and October 1958 reveals a positive relationship between short-time noise power decreases and solar flares. The most significant noise fadeouts associated with flare eruptions occurred with the sun over one of the major noise centers contributing to the noise level at the measuring station. Maximum noise decrease of 18 db was observed on 1 Mc/s at Kekaha when the sun was over the East Indies noise center, just after sunset at the river.

Introduction. Geographical regions having a high amount of thunderstorm activity, such as tropical Africa, South America, and the East Indies, act as large sources of radio noise that is received at distant points by ionospheric propagation modes. The received noise power is much the same as ordinary radio signals received from transmitters. Therefore it may be expected that ionospheric disturbances which affect radio transmissions will also affect atmospheric noise levels. The present investigation was undertaken to determine the effects of sudden ionospheric disturbance (SID) on atmospheric noise.

From the noise-measuring stations at Kekaha, Hawaii (22°N, 159.7°W), and Ohira, Japan (35.6°N, 140.5°E), were selected for study because these two locations (see Fig. 1) are at places where short-wave fadeouts (SWF) associated with solar flares are pronounced. These stations are in the CRPL noise-measuring network and use the National Bureau of Standards radio noise recorder model ARN-2 (Low, Samson, Disney, and Jenkins, 1959). Frequency is monitored for 15 minutes of each hour, and the average noise level during this period is assumed to be representative of the full hour. The noise-recording times were 0000-0100 hours and 2115-2200 hours for 2.5 and 5.0 Mc/s, respectively. The data used in this paper are such 'hourly' values.

The major noise sources for Kekaha are those in the East Indies (about 100-115°E longitude) and South America (about 60-75°W longitude).

The major source for Ohira is in the East Indies, two secondary sources being located in South America and equatorial Africa (about 0-15°E longitude). The noise output of these sources is not constant, but varies with time of day and season. In Table 1 the diurnal variations in activity of the sources for the months considered here are summarized. The measure of activity is taken to be the average noise power predicted for 1.0 Mc/s in 4-hour (GMT) time blocks in the region of the noise centers [International Radio Consultative Committee, 1957].

As a large number of solar flares occurred during August, September, and October 1958, the radio noise data for these months on the frequencies of 2.5 and 5.0 Mc/s were investigated. These particular frequencies were chosen because of their wide usage in ground-to-aircraft communications.

Method of analysis. All observing stations reporting short-wave fadeouts of importance 2 or greater [CRPL, 1958] during solar flares were plotted on maps as in Figure 1. The solid dots mark the SWF reporting stations, and the circled dots identify the two noise-recording stations. By plotting the SWF in this manner, an estimate of the widespread effect of the solar flare is obtained. Also, the overhead position of the sun with respect to noise sources and recording stations is easily visualized by letting the 0° longitudinal line represent 1200 GMT. For example, during the event plotted in Figure 1 (Sept. 15, 1958, 1700-1750 GMT) the sun moves from 75° to about 90°W longitude, effectively blanketing the noise source in South America.

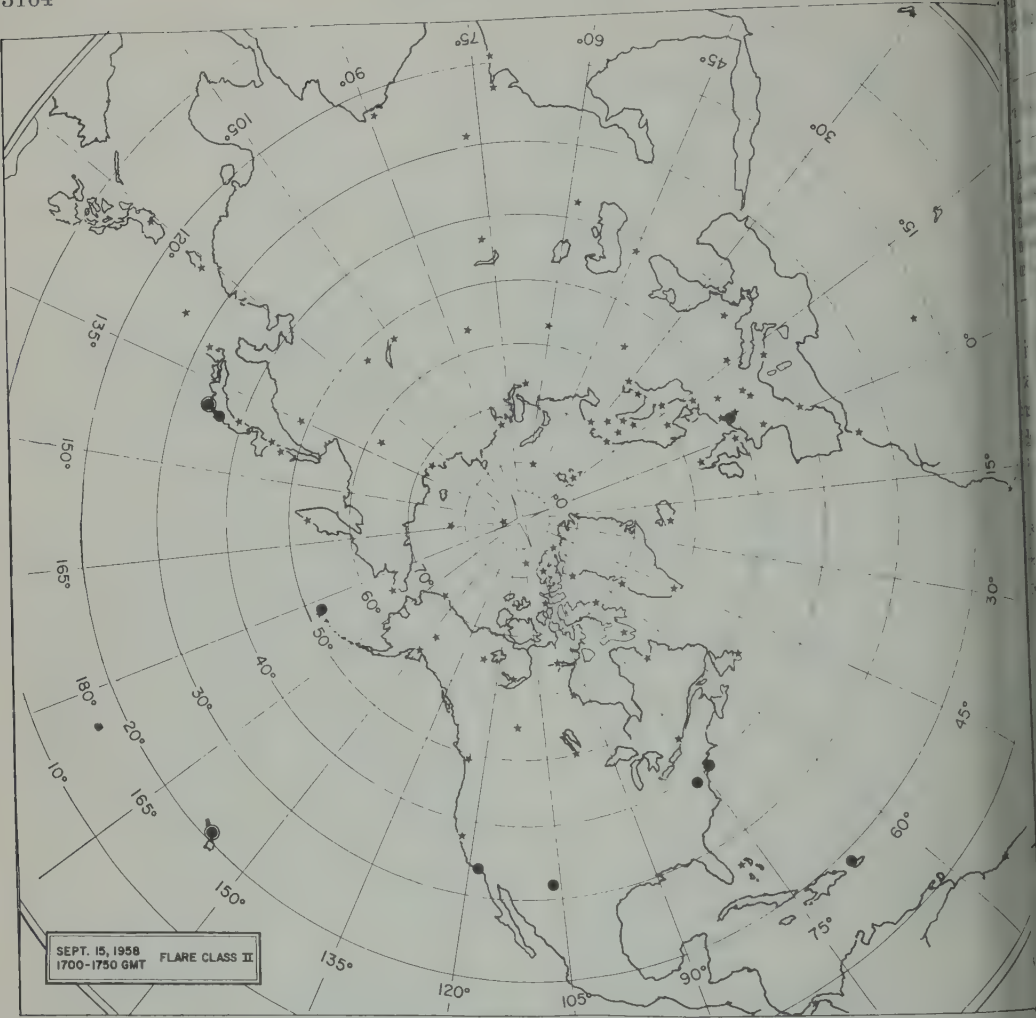


Fig. 1. Location of radio-noise-measuring stations (circled dots) and stations observing short-wave fadeout (solid dots).

Table 2 gives the estimated time intervals in which the listed geographical regions are most affected by the occurrence of a solar flare.

The diurnal behavior of the 2.5 and 5.0 Mc/s atmospheric noise is very regular, being high at night and low during the day, as is evident from the median curves in Figures 2 and 3. This shows that at night a substantial portion of the received noise power is propagated via the ionosphere from distant sources, which is absorbed in the ordinary *D* region during the day. The daytime level thus consists mostly of locally generated noise, and the occurrence of a solar flare would affect the noise level on these two frequencies

TABLE 1. Maximum Activity of Major Radio Noise Centers

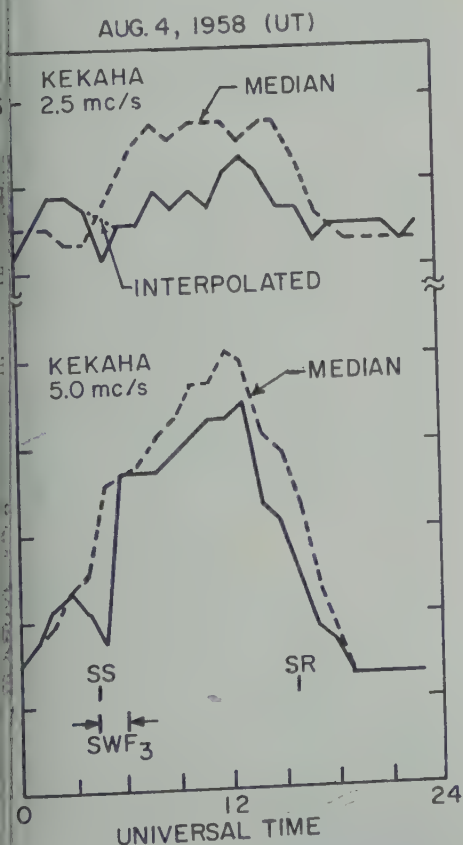
Levels listed are noise power in decibels above *kTB* predicted for 1.0 Mc/s in the region of the centers. All times are GMT.

Time Block	South America	East Indies	Equatorial Africa
0000-0400	100	55	100
0400-0800	100	60	80
0800-1200	75	80	75
1200-1600	70	85	70
1600-2000	75	90	90
2000-2400	90	70	100

2. Time Intervals in Which Listed Geo-Regions Are Most Affected by Occurrence of Solar Flare

Most Affected	GMT Time Interval
America source	1400-2000
Noise station	2000-0200
Indies source	0200-0800
Noise station	2400-0600
Equatorial Africa Source	0900-1500

at all. Therefore, no solar-flare effect is during the periods 2000-0200 GMT at and 2400-0600 GMT at Ohira. re occurring in the 1400-2000 interval effectively blanket the noise source in America, with a corresponding decrease in power recorded at Kekaha. However, a dropout to local noise level would not be expected because some sky-wave contribution



2. Radio noise recorded at Kekaha, August 4, 1958. Sunset and sunrise indicated by SS and SR.

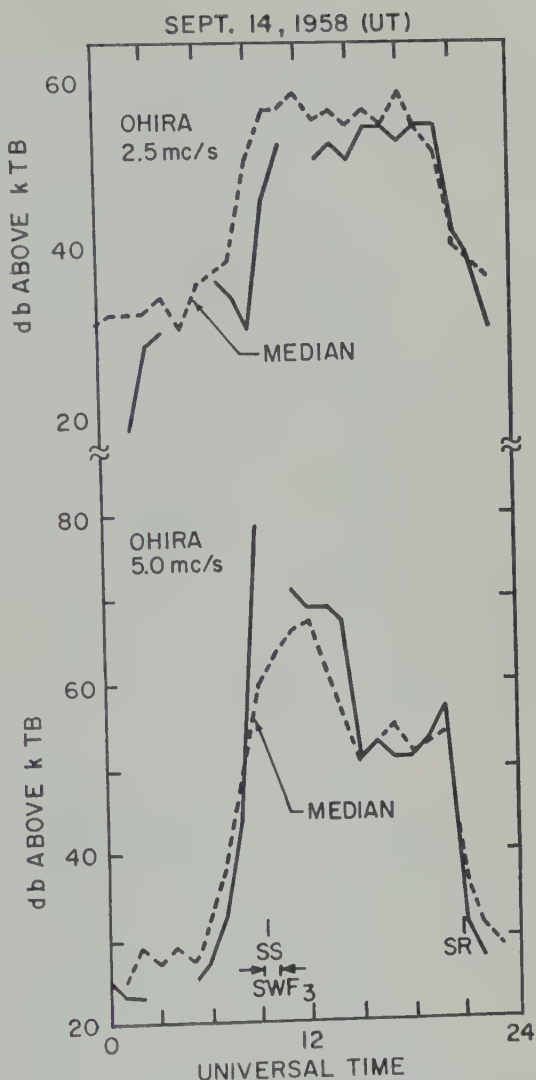


Fig. 3. Radio noise recorded at Ohira, September 14, 1958. Sunset and sunrise indicated by SS and SR.

would be received from the East Indies. For the interval 0200-0800 a similar situation would exist, with the roles of the East Indies and South American sources reversed. Between 0800 and 1400 no noise fadeout would be expected at Kekaha.

Similar reasoning shows the relation between Ohira, the East Indies, and equatorial Africa. In addition, according to the scheme in Table 2 both of Ohira's secondary sources (equatorial Africa and South America) would be blanketed between 1400 and 1500. At this time the East

TABLE 3. Magnitude of Noise Power Decreases at Kekaha during Solar Flares
Symbol N indicates record obscured by local noise; C means equipment failure.

Date 1958	GMT	Noise Power Decrease, db		Solar Flare Im- portance
		2.5 Mc/s	5.0 Mc/s	
Aug. 13	0113-0209	6	0	2
Oct. 18	0117-0150	0	6	2
Oct. 14	0249-0326	0	7	1
Aug. 9	0350-0430	8	12	2+
Sept. 18	0400-0553	2	2	2
Aug. 4	0422-0558	6	18	3
Aug. 16	0432-0720	6	6	3+
Sept. 12	0700-0742	5	3	2+
Aug. 18	0806-0900	6	0	2
Aug. 14	1613-1705	0	0	2
Sept. 7	1658-1743	10	C	2
Sept. 15	1700-1750	2	7	2
Sept. 2	1700-1732	C	C	1+
Oct. 17	1709-1853	N	N	1+
Oct. 4	1756-1823	4	0	1+
Oct. 13	1920-2005	1	0	2+
Sept. 25	2253-2332	C	4	1

Indies source is most active of the three (Table 1), and, since this source as well as Ohira are both in darkness, a flare occurring during this period would not cause a significant noise power decrease at Ohira.

In determining solar-flare effects on the radio noise the plotted SWF maps were used to establish which events might have caused noise fadeouts, and then the noise power data were checked during these periods. Solar-flare events occurring between 0800 and 1400 were excluded from the Kekaha analysis, as were those occurring in time intervals when the noise level was not being recorded. The noise records contain some short-time 'fadeout' depressions similar in appearance to those under discussion, but not accompanied by a reported solar flare or short-wave fadeout. No attempt is made in this paper to explain these variations.

Results. For illustration, the noise power recorded at Kekaha on August 4, 1958, is shown in Figure 2. The dashed curve is the August median noise level. During the reported SWF, 0422-0558 GMT, it can be seen that a definite noise power decrease was recorded even though the event took place just after ground sunset at

Kekaha. Figure 3 shows the September 14, 1958, Ohira noise data and monthly median. The noise power dropout is marked on the 2.5 Mc/s curve, but local interference obscures the 5.0 Mc/s record.

The noise power curves shown in Figures 2 and 3 are fairly typical of the data recorded near (receiver) local sunset. The 5.0 Mc/s Kekaha data in Figure 2 and the 2.5 Mc/s Ohira data in Figure 3 are quite similar during the SWF intervals. However, the noise fadeouts occurring just before (receiver) local sunrise are more characterized by the sharp dropout and recovery typical in Figures 2 and 3. The dropout is sharp but the recovery is usually masked by ordinary daytime absorption on the frequencies considered here.

The 2.5 Mc/s data for August 4 at Kekaha (Fig. 2) are interesting. A noise fadeout accompanies the SWF between 0422 and 0558 GMT, but the usual nighttime level (as depicted by the median curve) is not attained afterward. This low nighttime level in early August has been attributed to the effects of the Johnston Island nuclear explosions [Samson, 1960].

The results of the Kekaha analysis are given in Table 3, and those for Ohira in Table 4. Times of the events, coinciding with the reported SWF times, are listed in increasing order of GMT time, as the dates have no particular significance. The importance of the associated solar flare is shown in the last column. The magnitude of the noise power decrease is shown in the preceding columns.

TABLE 4. Magnitude of Noise Power Decreases at Ohira during Solar Flares

Symbols N and I indicate record obscured by local noise and radio interference, respectively.

Date 1958	GMT Time	Noise Power Decrease, db		Solar Flare Im- portance
		2.5 Mc/s	5.0 Mc/s	
Aug. 13	0113-0209	2	2	2
Aug. 9	0350-0430	0	0	2+
Aug. 12	0421-0553	N	N	2
Aug. 16	0432-0720	0	14	3+
Sept. 4	0507-0554	0	2	1
Sept. 12	0700-0742	N	N	2+
Aug. 18	0806-0900	8	5	2
Sept. 14	0851-0949	18	I	2+
Sept. 2	2105-2137	0	6	1

bels below the interpolated noise level the flare not occurred. The method of interpolation is illustrated in Figure 2 on the μ/s curve. The solid curve connects the measured values, and the dotted line is the noise level interpolated over the flare period. In the cases, as for the 5.0 Mc/s record shown in Figure 2, the recorded noise level coincided with the monthly median before and after the flare event but was depressed during the

flare. In the data in Table 3 it is seen that the atmospheric radio noise power measured at Kekaha is most affected by the occurrence of a solar flare when one of the two major noise centers in the East Indies or South America is daylit. No effect was observed when the sun was within an hour and a half on either side of maximum elevation at Kekaha, the period of ordinary D -region ionization absorbs most of the 2.5 and 5.0 Mc/s noise signals. Also no effect was observed at Kekaha when the solar flare occurred between 1400 and 1600 GMT, indicating that the enhanced D -region absorption associated with solar flare was not intense enough to cut off sky-wave-propagation of noise to Kekaha.

The maximum noise power decrease (18 db on 5.0 Mc/s) at Kekaha was observed between 0600 and 0600 GMT, during which time the sun was overhead over the major noise center in the East Indies. The corresponding time interval for the sun over the noise center in South America is 1600–1800 GMT, and during this time the maximum noise power decrease observed at Kekaha was 7 db on 5.0 Mc/s. In this comparison it appears that the East Indies source contributes more to Kekaha's noise level than the source in South America.

Noise power decreases at Ohira during solar flares are listed in Table 4. Nine events in which the flare might have affected the noise level have been omitted from the compilation because of equipment failure at Ohira.

The limited amount of data in Table 4 indicates that the Ohira noise power decreases are most pronounced when the solar flare occurs between 0600 and 1000 GMT. During these hours the sun is daylit over the noise centers in both the East Indies and equatorial Africa, so that the enhanced D -region absorption accompanying the flare would reduce

the intensity of the noise signals propagated to Ohira. The 14-db decrease on 5.0 Mc/s at Ohira (Aug. 16, 0432–0720) was recorded at 0700; the noise fadeout was less severe for the 0500–0600 measurements. It is notable that the largest observed noise power decreases occurred near ground sunset at Ohira.

Conclusions. Although about 75 flares accompanied by SWF occurred during August, September, and October 1958 which fit the acceptability criteria of the present analysis, a large number of events had to be rejected because of equipment failure or because the event happened entirely during recorder 'off time.' Nevertheless, the data given here are considered sufficient to establish a positive relationship between solar flares and atmospheric radio noise power decreases.

For the two frequencies of 2.5 and 5.0 Mc/s, it appears that solar-flare effects are most pronounced when one of the major noise centers is blanketed by a sudden ionospheric disturbance. Near local noon at the noise-measuring stations ordinary D -region absorption effectively blocks reception of these frequencies by sky-wave modes, so that no discernible noise power change accompanies a flare eruption.

Acknowledgments. The work described in this paper has been supported by Air Force Cambridge Research Laboratories under contract AF19(604)-4092. Dr. W. Q. Crichlow of Central Radio Propagation Laboratory kindly supplied the hourly noise data from Kekaha and Ohira. I am indebted to Mr. G. E. Hill for the method of approach used in this analysis.

REFERENCES

- Central Radio Propagation Laboratory, *Solar-Geophysical Data, CRPL-F 170*, part B (serial), 1958.
- Crichlow, W. Q., C. A. Samson, R. T. Disney, and M. A. Jenkins, Radio noise data for the International Geophysical Year July 1, 1957–December 31, 1958, U. S. Department of Commerce, *NBS Tech. Note 18*, 1959.
- International Radio Consultative Committee, Revision of atmospheric radio noise data, *C.C.I.R. Rept., 65*, International Telecommunication Union, Geneva, 1957.
- Samson, C. A., Effects of high-altitude nuclear explosions on radio noise, *J. Research NBS*, **64D**, 37–40, 1960.

(Manuscript received June 20, 1961.)

Infrared and Reflected Solar Radiation Measurements from the Tiros II Meteorological Satellite

R. R. BANDEEN, R. A. HANEL, JOHN LIGHT,¹ R. A. STAMPFL, AND W. G. STROUD

*Goddard Space Flight Center
National Aeronautics and Space Administration
Greenbelt, Maryland*

Abstract. The Tiros meteorological satellite contains detectors, storage, and telemetry for the measurement of infrared and reflected solar radiation from the earth and its atmosphere. Two separate detector designs are employed: a medium-resolution scanning radiometer and a low-resolution scanning radiometer. The spin of the satellite provides the scan line of the medium-resolution scanning radiometer, which is then advanced by the orbital motion of the satellite. Five channels using radiometer detectors and filters to limit the spectral response from 6 to 6.5 microns, 8 to 12 microns, 5 to 6 microns, 8 to 30 microns, and 0.55 to 0.75 micron are mounted in a single housing with amplifiers and pre-amplifiers. The spatial resolution is about 40 miles square when viewing the earth directly beneath the satellite. The parameters studied by these spectral regions are, in the order: radiation emerging in the water vapor absorption band, day and night time cloud cover, albedo, thermal radiation, and visual maps for comparison with television pictures from vidicon cameras also carried in the satellite. The low-resolution nonscanning radiometer, using a simple unchopped design, measures the blackbody temperature and the albedo of the earth. The field of view of the detector when viewing the earth directly beneath the satellite is a circle of 450-mile diameter and covers part of the area of each picture frame of the wide-field television camera. The detector consists of two thermistors, each mounted in the apex of a reflective cone which provides optical gain. One thermistor, coated black, responds to both reflected solar radiation and the thermal radiation from the earth; the second reflects solar radiation and responds only to the thermal radiation. The design, calibration, performance, and data reduction for both systems are discussed.

Introduction. The Tiros II meteorological satellite was launched into orbit on November 23, 1960 (Fig. 1). Its orbital characteristics are listed in Table 1 together with those of the first Tiros for comparison. Like its predecessor, Tiros I, it carries two television cameras. Although pictures from the wide-angle camera are not transmitted to their Tiros I counterparts in orbit, the Tiros II television system has worked with considerable success. The functioning of the television system, which is practically identical to that of Tiros I, has been reported previously and will not be discussed here [Stroud, Stemberg and Stroud, 1960].

The outstanding difference between the two satellites was the inclusion in Tiros II of an experiment consisting of two radiometers to measure the infrared and reflected solar radiation from the earth and its atmosphere. It is on this experiment that we shall report in this paper.

Work with Lockheed Electronics Corp., Sunnyvale, New Jersey.

Description of the experiment. The Tiros II radiation experiment consists of a medium-resolution five-channel scanning radiometer whose bidirectional optical axes are inclined to the spin axis by 45° and 135° and of a low-resolution two-channel nonscanning radiometer whose optical axes are parallel to the spin axis. As Tiros spins, the 5° field of view of each medium-resolution radiometer channel scans the earth and outer space alternately. The orbital motion of the satellite provides the advancement from one sweep line to the next. The spin rate of about 12 rpm was chosen to accommodate a proper scan pattern, that is, a raster without over- or underlap of individual lines. The spin rate is consistent with the available information bandwidth of 8 cps per channel. A scan line is given by the intersection of a 45° half-angle cone described by the optical axis of the detectors in one spin cycle and the earth's near-spherical surface. The scan pattern is a circle on earth wherever the spin axis is parallel to the local earth radius vector, and it becomes a pair of

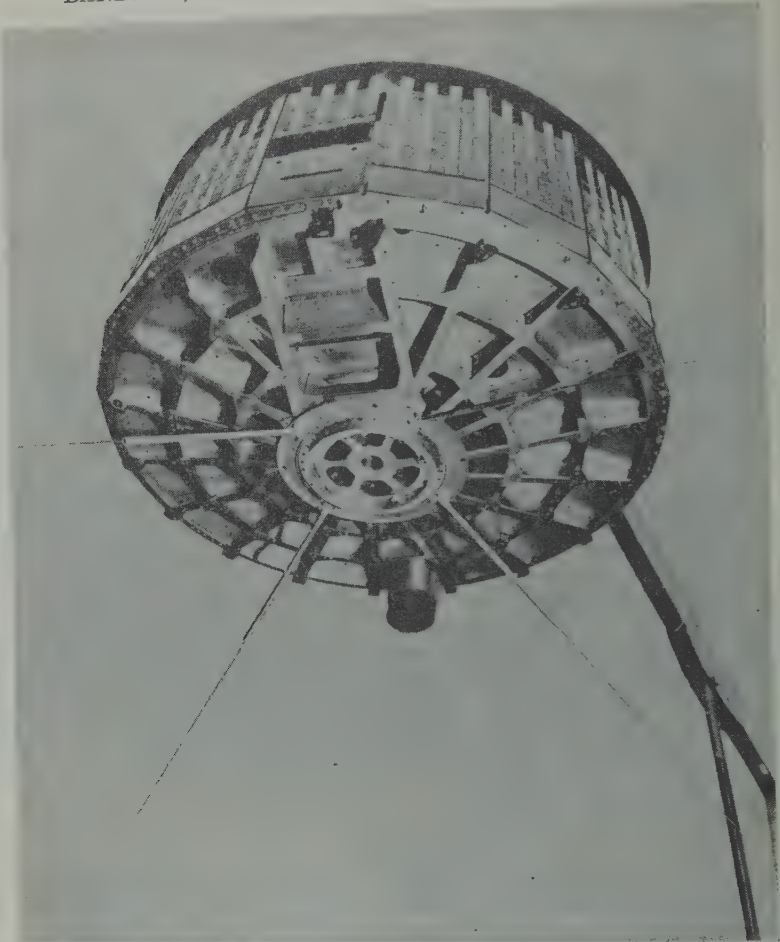


Fig. 1. Tiros II meteorological satellite launched November 23, 1960. The medium-resolution scanning radiometer looks through rectangular apertures in the side and base plate. The low-resolution radiometer looks through the round aperture in the base plate almost diametrically opposed to the protuding wide-angle television lens.

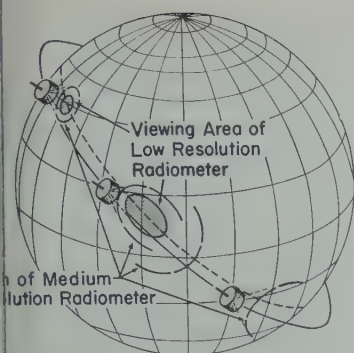
alternating, hyperbola-like branches wherever the spin vector is normal to the radius vector. The geometry of the motion is identical to the geometry the Vanguard II cloud-cover satellite was supposed to (but actually did not) assume

TABLE 1. Orbital Characteristics of Tiros I and II

	Tiros I	Tiros II
Date launched	Apr. 1, 1960	Nov. 23, 1960
Perigee, statute miles	432.9	385.6
Apogee, statute miles	464.4	454.4
Period, minutes	99.24	98.27
Inclination, degrees	48.39	48.53

[Hanel, Licht, Nordberg, Stampfl, and Stroud 1960]. The spin of the satellite, however, does not modulate the 50° field-of-view low-resolution radiometer channels because their optical axes are parallel to the spin axis. The low-resolution radiometer observes an area that is within the field of the wide-angle TV camera. Interesting correlations between cloud cover and heat balance are expected. The geometry of the scanning motion of the medium-resolution and of the viewing area of the low-resolution radiometers is shown in Figure 2.

The instrument in the satellite measures radiation intensities in a certain direction and from a certain area, namely, the area in the field of view



Geometry of the scanning motion of a medium-resolution radiometer and of the viewing area of a low-resolution radiometer.

scattering function of this area were known, the amount of backscattered energy, the total energy could be computed. However, because the earth is not in the field of view neither reflects nor emits according to Lambert's law (like a diffuse surface), assumptions about the isotropic nature of radiation must be made. The total radiation loss in all directions [Mark and Yamamoto, 1961]. Calculations of the anisotropic nature of radiation emerging from the atmosphere can be based on model atmospheres. Certain areas on the earth's surface are observed by the satellite under different zenith angles within a time interval of about 3 minutes. Results of these observations tend to verify the choice of a particular model atmosphere in a given region [Greenfield and Logg, 1960; Wexler, 1959]. The medium-resolution radiometer, lens

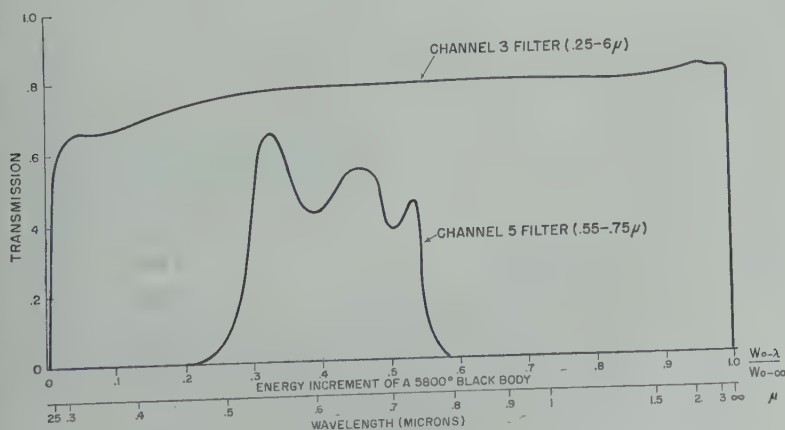
materials and filters restrict sensitivities of the five channels to the following spectral regions:

1. 6-6.5 microns, water vapor absorption.
2. 8-12 microns, atmospheric window.
3. 0.25-6 microns, reflected solar radiation.
4. 8-30 microns, thermal radiation.
5. 0.55-0.75 micron, visible reference and day-time cloud cover.

The physical significance of these regions has been reported previously [Hanel and Stroud, 1960] and will be discussed only briefly here.

In Figure 3 the transmission characteristics of the filters of channels 3 and 5 are plotted vs. the available energy from a blackbody of 5800°K, the color temperature of the sun. About 99 per cent of backscattered and reflected sunlight falls into the spectral range of channel 3. The spectral range of channel 5 was chosen to give good contrast between earth and clouds. It is in the range of visible and infrared photography and close to the spectral sensitivity of television cameras in Tiros. This channel yields cloud-cover pictures on the illuminated side of the earth, whereas the television system covers only limited areas, although with a much higher resolution.

The transmission characteristics of the three thermal channels are shown in Figure 4. The abscissa is linearly proportional to the energy available from a 300°K blackbody. The diagram does not include the emissivity of the thermistor bolometer or the chopper characteristic. The spectrum of channel 4 from 8 to 30 microns covers fairly well the range of thermal emission from our planet. For a study of the energy



3. Filter transmission characteristics of channels 3 and 5 of the medium-resolution radiometer.

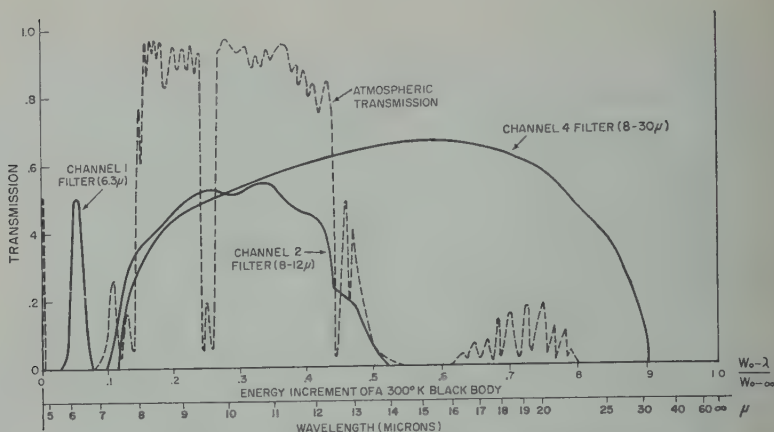


Fig. 4. Filter transmission characteristics of channels 1, 2, and 4 of the medium-resolution radiometer. The dashed line is the approximate transmission characteristic of 1 atmosphere.

budget the total amount of radiation loss is even more important than the albedo. Radiation emerging in the atmospheric window between 8 and 12 microns is measured by channel 2. The atmosphere is fairly transparent in this spectral range, and the apparent blackbody temperatures are close to the true temperatures of the radiation surfaces, whatever they may be—cloud tops, water, or land. Corrections for ozone absorption near 9.6 microns will be applied. Since clouds are generally cooler than the surface of the earth, a map showing isolines of radiant emittance can be interpreted as a cloud-cover map. This method is especially valuable since it works also on the dark side of the earth, which is unobserved by television cameras. The difference between channel 4 and channel 2 is essentially radiation between 12 and 30 microns, characterized by strong absorption bands of carbon dioxide and water vapor. The spectrum of channel 1 corresponds to the region of water vapor absorption between 6 and 6.5 microns. The temperature

profile and the relative humidity in the atmosphere determine the energy that can be observed by this channel.

The spin vector of the Tiros I satellite was observed to exhibit angular motions of large amplitude during its 78-day active life which were shown to compare closely with a theoretical model based on reactions to two external torques, namely, a primary torque caused by the interaction of a magnetic dipole along the satellite spin axis with the earth's magnetic field, and a secondary torque caused by differential gravitation in the earth's gravitational field [Bandeem and Manger, 1960]. Because of the importance of keeping the radiation sensors from viewing the sun longer than absolutely necessary, it was decided to include a closed current loop in Tiros II with several possible levels of current flow, both positive and negative, any one of which could be commanded from the ground. Thus the motion of the satellite spin axis could in some measure be controlled. The magnetic attitude control has worked satisfactorily, and observed spin axis motions after programming specific current values in the magnetic attitude control coil have agreed well with calculated theoretical movements.

Spin-up rockets attached to the periphery of the satellite base plate can be fired upon command from the ground to achieve the desired spin rate if the despun after injection results in too low a rate, or to restore the desired level after magnetic spin-decaying effects have operated for a period of time. Two pairs of spin-up rockets were

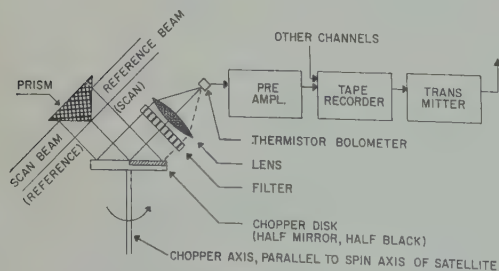


Fig. 5. Block diagram of one channel of the medium-resolution radiometer.

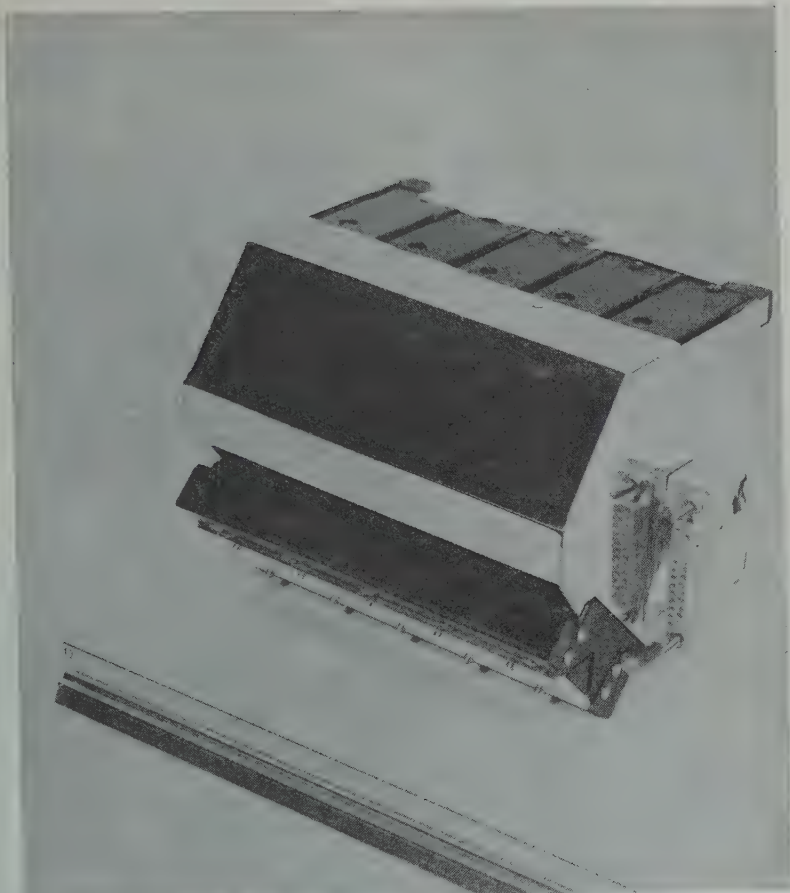


Fig. 6. Exterior view of the medium-resolution radiometer, showing the view apertures in one section of the five channels. The prismatic cross section of the reflector is seen on the right line of apertures.

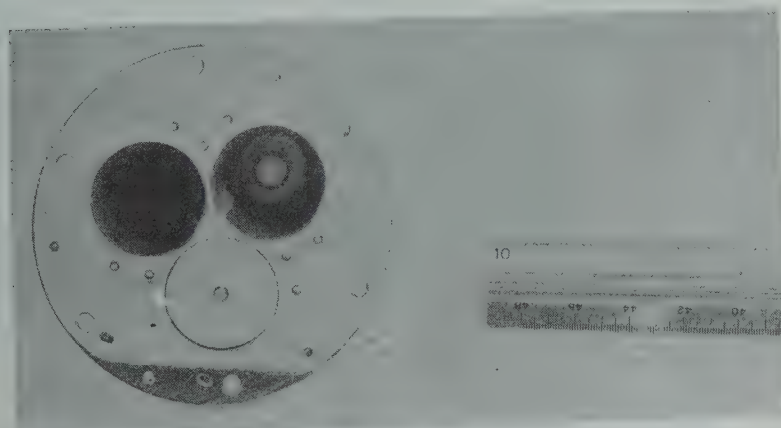


Fig. 7. Exterior view of the low-resolution radiometer showing the black detector (left) and the white detector (right).

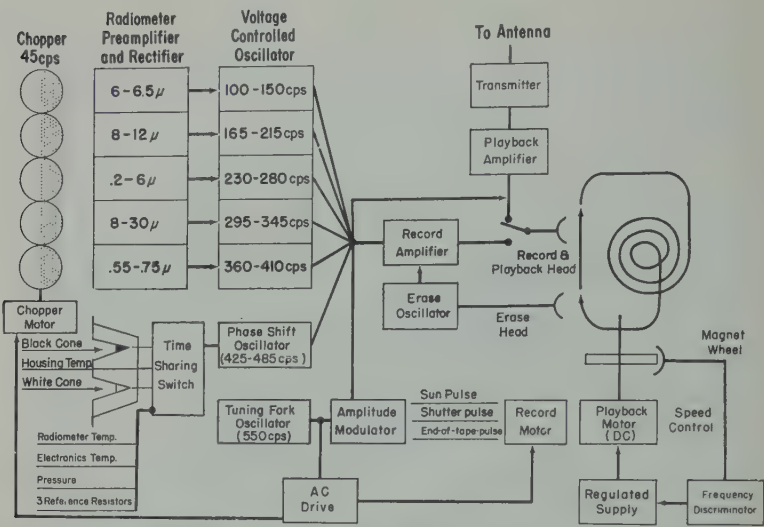


Fig. 8. Block diagram of the radiation experiment.

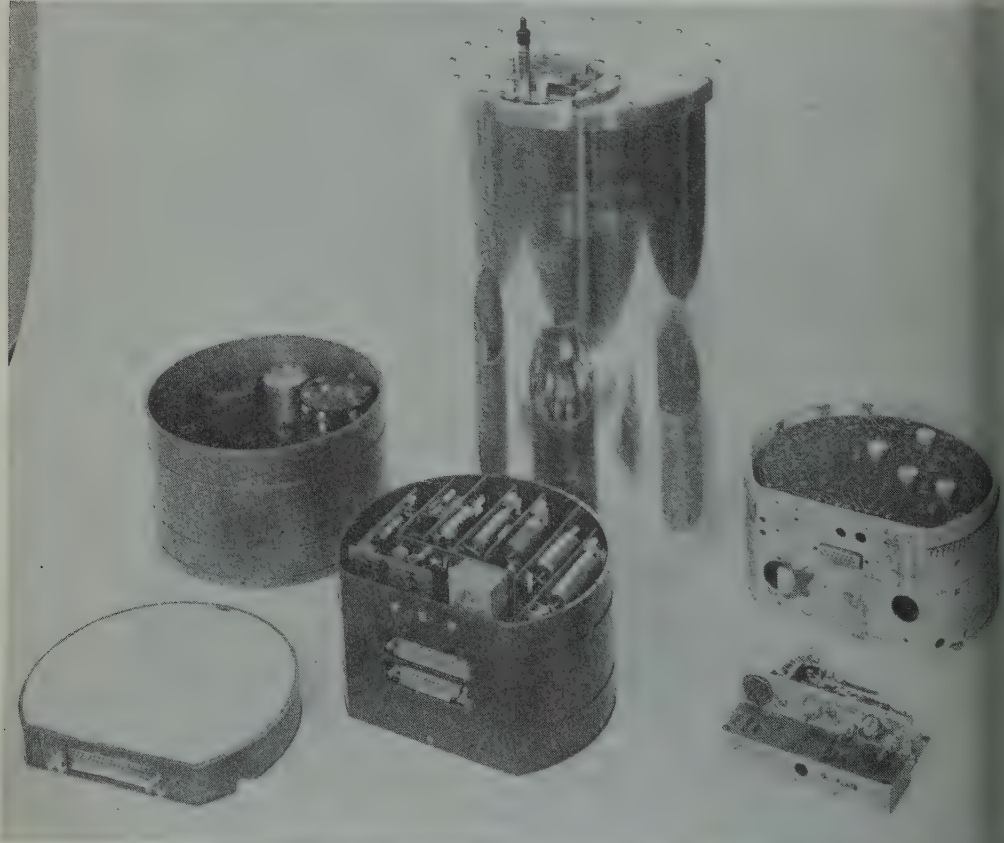
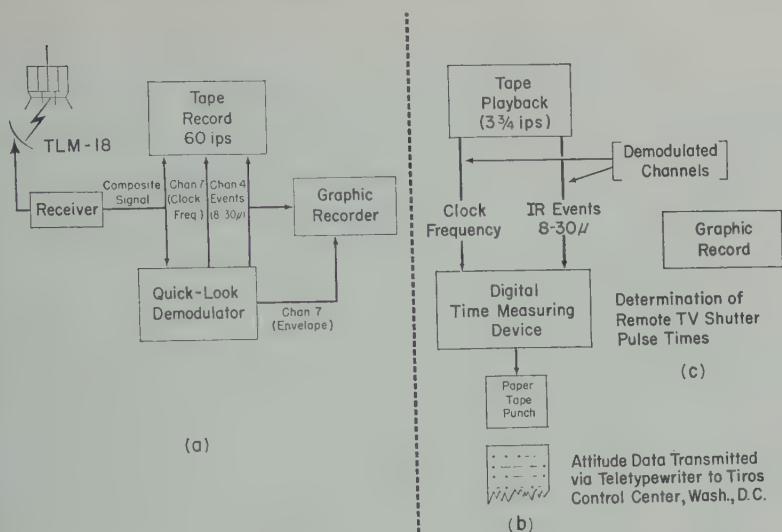


Fig. 9. Exploded view of the components of the radiation experiment canister. From left to right they are: power supply, tape recorder, main deck electronics, and the 2-watt transmitter. All components fit into the canister to the center rear.



10. Block diagram of information flow at a data acquisition station, including auxiliary uses of the radiation data.

to days after launch increasing the initial rate of 8.0 rpm to 14.0 rpm.

umentation. One channel of the medium-resolution radiometer is shown in Figure 5. The disk, half reflecting, half absorbing, totally reflects radiation from the scan beam to the reference beam to the detector. Consequently, the alternating voltage generated at

the thermistor bolometer is proportional to the energy difference between the two opposite directions. The radiometer is shown in Figure 6. The two low-resolution radiometer channels consist of a black and white detector each mounted in the apex of a highly reflective cone (Fig. 7) [Hanel, 1961]. The black detector is equally sensitive to reflected sunlight and long-

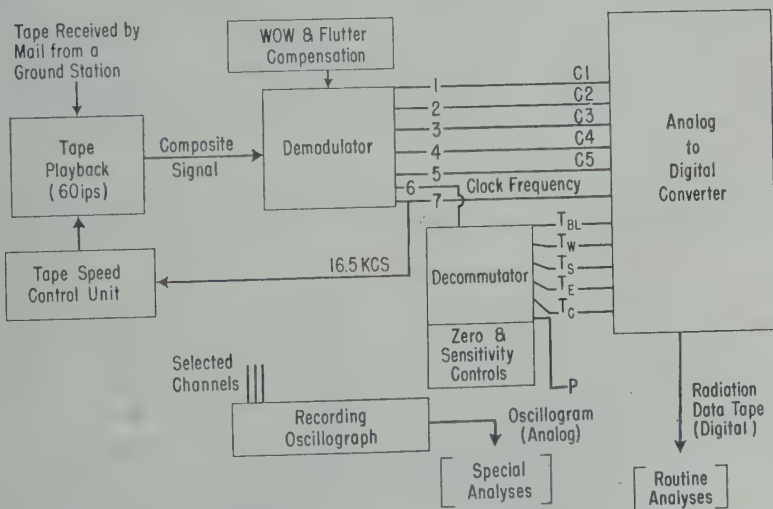


Fig. 11. Block diagram of information flow at the data reduction center in producing a digital magnetic tape for computer input. The output of the decommutator consists of temperatures of the black and white cones and three environmental temperatures and canister pressure.

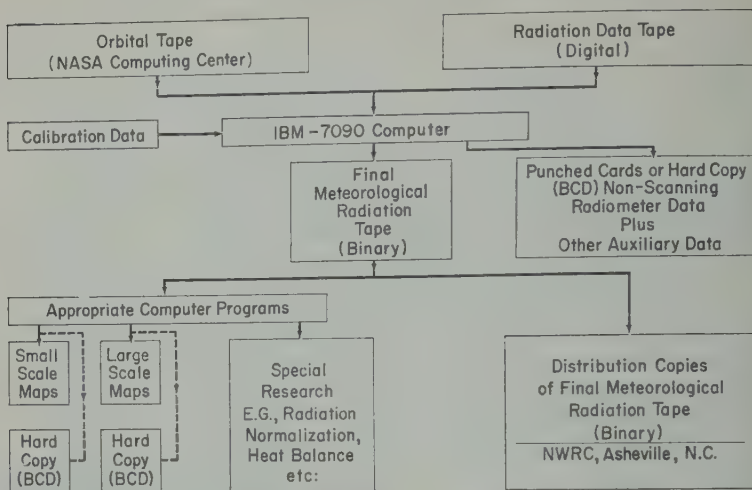


Fig. 12. Block diagram of information flow through the IBM-7090 computer at the data reduction center.

wave terrestrial radiation. The white detector is coated to be reflective in the visible and near infrared. The surface appears white to the eye even though its emissivity is high in the far infrared. Since 99.9 per cent of the terrestrial radiation is emitted at wavelengths of 4 microns and longer, both detectors show the same equilibrium temperature when they face the dark side of the earth. On the illuminated side the temperature of the black detector will rise in contrast to that of the white one, which is affected to a lesser extent by reflected sunlight. Careful measurement of the temperature of the detectors and the mounting plate, as well as determination of thermal conduction and radiation coupling between thermistor flake and satellite, allows us to determine the apparent blackbody temperature and the amount of reflected sunlight within the 50° field of view.

Before installation in the satellite the radiometers had to pass a series of environmental tests and calibration measurements. To calibrate the three thermal channels of the medium-resolution radiometer, the radiometer was exposed to two blackbodies. One, simulating outer space, was kept at liquid nitrogen temperatures. The second blackbody, which simulated radiation from earth, was adjusted to various temperatures between 250° and 320°K . To minimize errors from water vapor absorption and to prevent condensation on the cold targets, the radiometer and the blackbodies were placed

in a dry nitrogen atmosphere. The response of a channel is proportional to the energy difference between opposite directions, expressed mathematically by

$$V = k \int_0^\infty [W_\lambda(T_1) - W_\lambda(T_2)] f_\lambda d\lambda \quad (1)$$

In equation 1, $W_\lambda(T)$ is the spectral radiance of a blackbody, and f_λ the filter function for a particular channel. Emissivity of the bolometer and reflectivity of the chopper have been taken as independent of wavelength. This is fairly well justified for all but the 8-9 micron channel. The second part of the integrand in equation 1 is negligible for liquid nitrogen temperatures, but it has to be taken into account if T_2 and T_1 are at room temperature and slightly above. The latter temperatures exist in a check of calibration at Cape Canaveral where the whole system including radiometer and telemetering electronics was tested. The constant k includes also the gain of the pre-amplifier. In the calibration the gain was set to give the same maximum voltage in each channel for different blackbody temperatures as follows: channel 1, 265°K ; channel 2, 315°K ; channel 3, 305°K .

The standard source in the calibration of the two solar channels of the medium-resolution radiometer (0.25-6 and 0.55-0.75 micron) was a tungsten band lamp with a quartz window. Extensive corrections for the lower color tem-

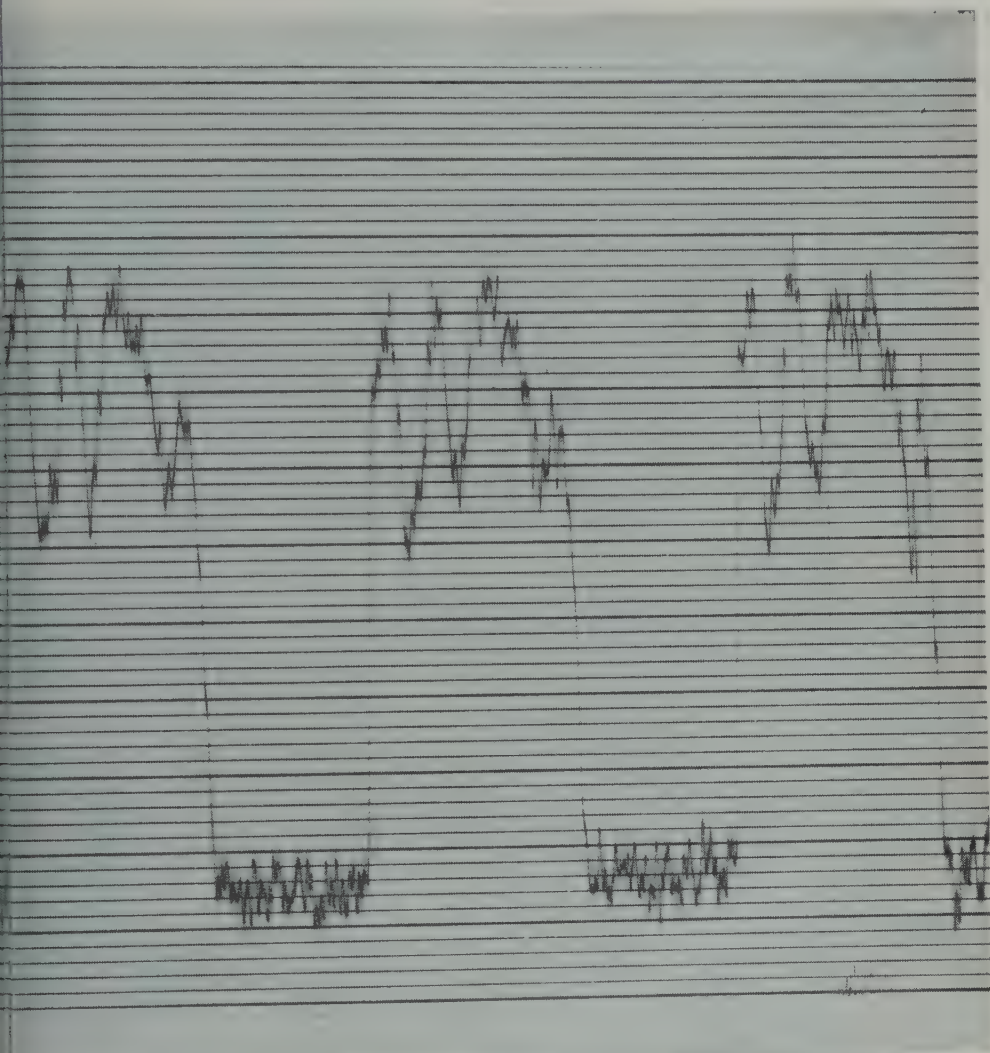


Fig. 13. Oscillogram showing three consecutive sky-earth scans of channel 2 of the medium-resolution radiometer. The spin period was 7.53 seconds. The amplitude is approximately proportional to the radiant energy received. Time increases to the right along the abscissa.

of 2200°K compared with sun temperature of about 5800°K were necessary. Since standard lamp is too small to fill the field of the radiometer (and the intensity would be too strong), another, larger source was an intermediate step. The opal glass secondary standard is not quite as good at 2 microns as it is in the visible. Unfortunately this effect was not recognized immediately, and the gain on these two channels was about 10 db too low. However, since the signal-to-noise ratio is good we can compensate

for this error in readjusting the gain in our ground station.

The low-resolution radiometer was calibrated in vacuum. The radiometer was mounted on a metal frame which simulated the satellite. The temperature of the frame was adjusted in steps between -10° and +60°C. Radiation from the earth was simulated by a blackbody which filled the field of view. The temperature of the blackbody was varied between -130° and +60°C. In this way, heat conduction and radiation coupling between thermistor flakes and satellite

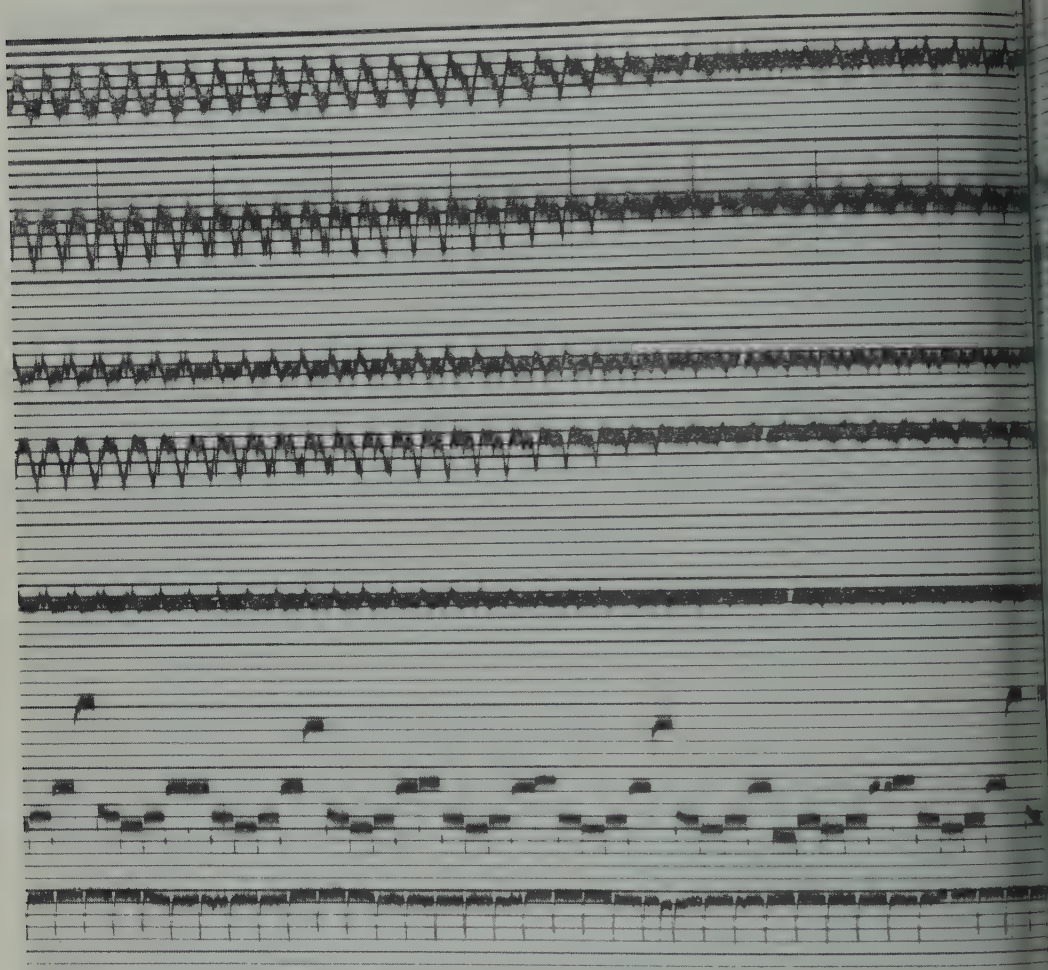


Fig. 14. Oscillogram showing all channels of the radiometer experiment. Reading from top to bottom: the five medium-resolution channels in numerical order, the commutated channel 6, and the envelope of the clock frequency showing sun sensor pulses every spin period of 7.53 seconds. The 'point of verticality' where one detector sweeps a circle on the earth can be recognized on the right where the horizon is not intercepted at all (see Fig. 2).

structure were determined. The detector, still in vacuum, was then exposed to sunlight (and artificial light) reflected by a white diffuser. The transmission characteristic of the quartz window that sealed the vacuum chamber was taken into account. A thermopile served as the calibration standard.

The radiation experiment instrumentation is independent of the television camera system except for power, command, certain timing

signals, and antennas (Figs. 8 and 9). The output of the five narrow-angle radiometer channels is fed to five subcarrier oscillators. These voltage-controlled oscillators are of the phase-shift type with symmetric amplifiers in the feedback loops the gains of which are controlled by the balance input signal. A sixth channel is provided for telemetry of the wide-angle low-resolution sensor data, environmental temperatures, instrumentation canister pressure, and calibration. A mecha-

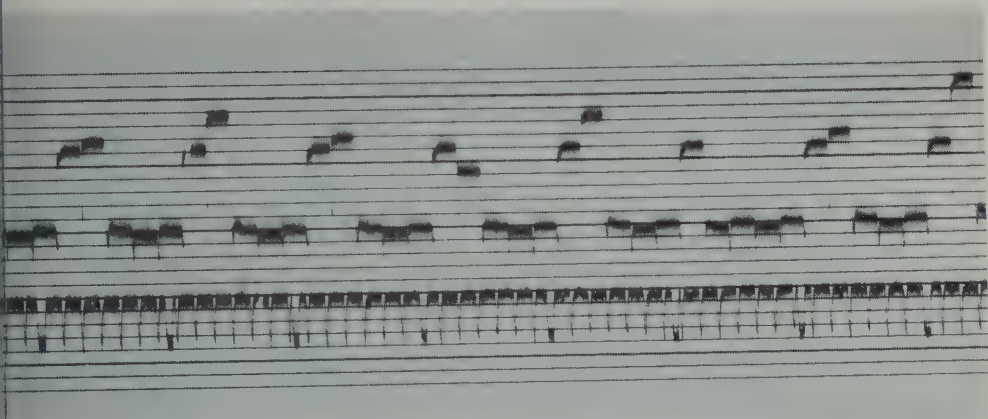


Fig. 15. Oscillogram of the radiation experiment, showing the last two channels. The wider pulses on channel 7 are camera shutter signals as pictures were taken every 30 seconds. The spin period here is 4.43 seconds resulting from firing two pairs of spin-up rockets on November 25, 1960.

commutator switches resistive sensors in one of a phase-shift oscillator. The seventh channel, a tuning-fork oscillator, serves as a reference frequency and timing signal. The signals from these seven different channels are combined, and the resultant composite signal is fed into a record amplifier which drives the motor of a miniature tape recorder. An oscillator provides an alternating-current bias to the record head and the signal required for the erase head. For convenience, erase of the magnetic tape occurs immediately before recording. The record rate extends from 100 to 550 cps. The tape recorder is an endless-loop, two-speed design operating at 0.4 ips record and 12 ips playback. The endless loop records continuously, day and night, except during a playback sequence. A synchronous motor generates torque for record mode through a Mylar belt speed reducer [Licht and White, 1960]. The fourth channel, a harmonic of the tuning-fork oscillator generated by flip flops drives the motor. The record motor also drives a camshaft which activates a set of microswitches connected to the five multiplexed subchannels of the time-sharing multiplex channel. Each is sampled 6 seconds, and each includes a group of seven to be submultiplexed.

Playback is initiated upon command by providing power to a direct-current motor. A balanced flywheel generates a frequency reference signal to the motor speed. A frequency

discriminator feeds the error signal to the stabilized power supply of the motor and closes the servo loop. Playback speed is essentially constant from 0° to 50°C. A low flutter and wow of 2.5 per cent peak-to-peak measured without frequency limitations is achieved by means of precision bearings and ground-in-place shafts having tolerances of better than 50 parts per million. A command pulse activates the playback motor, the playback amplifier, and the 238 Mc/s FM telemetry transmitter feeding the duplexers and antenna.

To permit comparison of the low-resolution measurements with TV pictures each TV shutter action generates a 1.5-second pulse which is recorded as an amplitude modulation of the channel 7 timing signal. As will be recalled from Tiros I, nine solar cells are mounted behind narrow slits for north angle determination. These slits have an opening angle close to 180° in planes through the spin axis. The sun illumination generates pulses as long as illumination parallel to the spin axis is avoided. One of these sensors generates a 0.5-second pulse in addition to the north indicator code so that spin rate information and a measure of relative sun position is available. Again, this pulse is recorded as an amplitude modulation of channel 7. Reconstruction of the radiation information vitally depends on its correlation with absolute time. An accurate but relative timing signal is provided by the tuning-fork oscillator and a

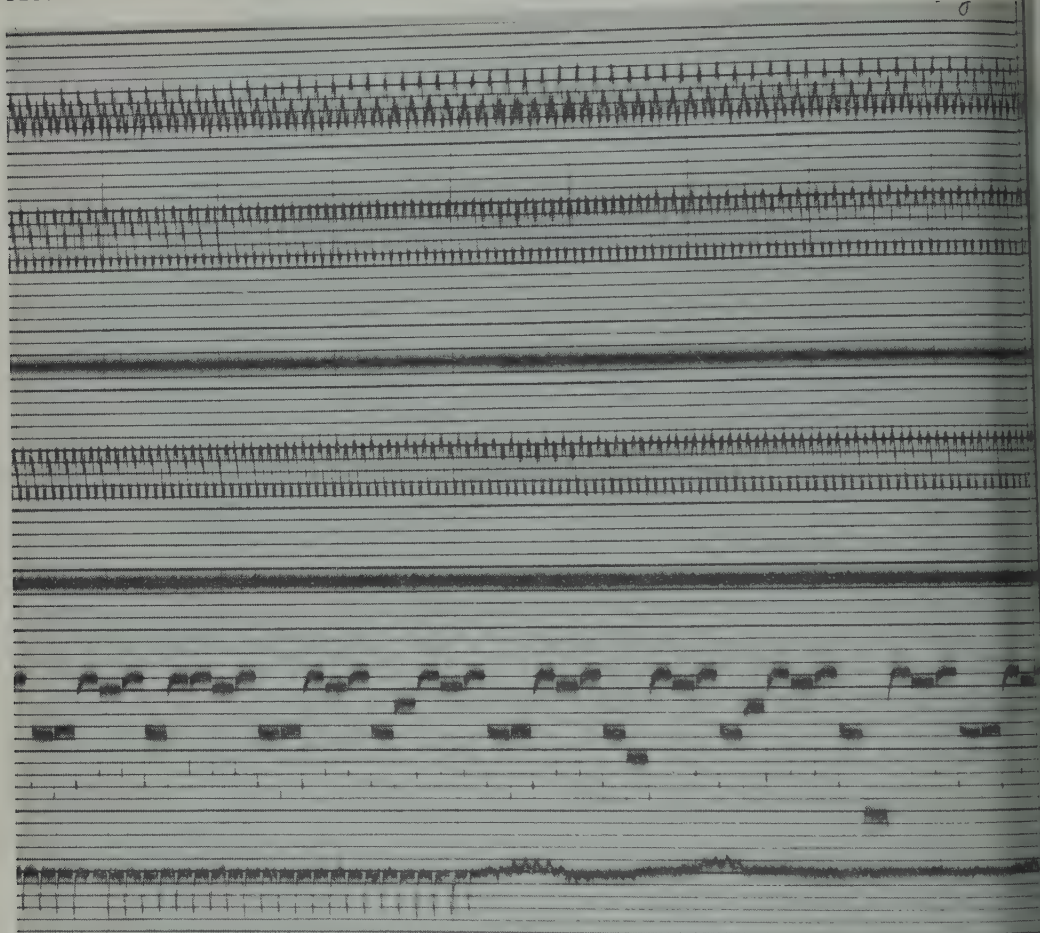


Fig. 16. Oscillogram of the radiation experiment showing passage into the earth's shadow, identified by the cessation of sun pulses in the right half of channel 7. The geometry was such that both directions of the medium-resolution radiometer alternately viewed the earth in one spin period (see Fig. 2). The spin period is 4.43 seconds.

crude one by the sun pulses except in the earth's umbra. Absolute time is transmitted to the satellite and recorded on the tape as a 1-second drop out of channel 7. The occurrence of this pulse is known within milliseconds of absolute time.

Operation of the experiment. Upon interrogation, the 238 Mc/s carrier is received by a 60-foot parabolic antenna, and the composite signal is recorded on magnetic tape and, simultaneously, fed to a 'quick-look' demodulator (Fig. 10a). At the same time, the envelope of channel 7 and the clipped signal of channel 4 are graphically recorded. The 8-30 micron 'events' on the graphic

record show alternately the earth and sky scan intervals as the satellite spins and progresses along the orbit, and the channel 7 envelope shows the three distinctive types of AM pulses impressed on the clock frequency during the record mode: the sun sensor pulses, the TV camera pulses, and the 'end-of-tape' pulse. Auxiliary uses of the radiation data include determination of the spin axis attitude in space and the times when television pictures were taken and recorded in the satellite, to be read out later over a ground station (Fig. 10b and c).

The magnetic tapes are routinely mailed every day to the Meteorology Branch, GSFC, i

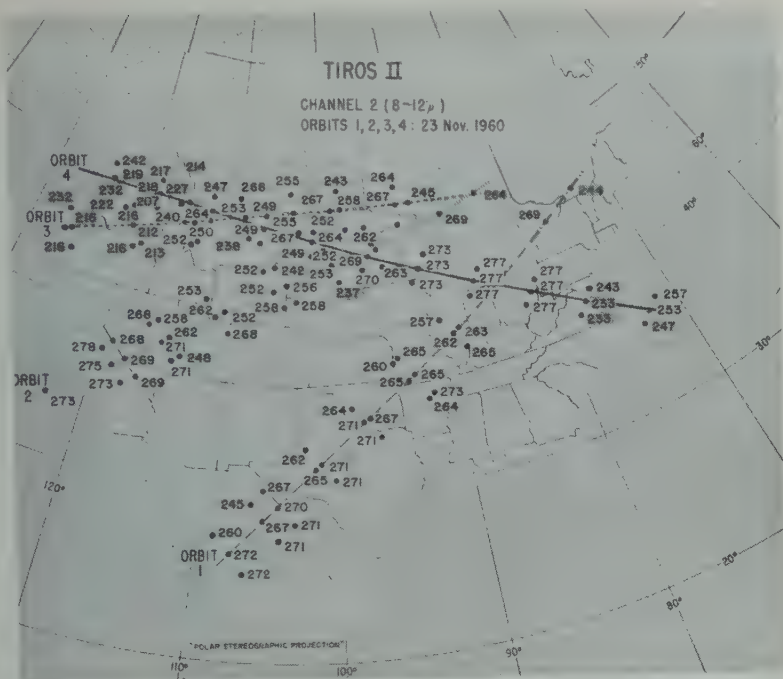


Fig. 17. Apparent blackbody temperatures viewed by the 8-12 micron channel of the medium-resolution radiometer while passing over the United States during the first four orbits after launch, November 23, 1960.

ington. The master tape containing the site radiation signal is demultiplexed, translated, and fed to an analog-to-digital converter (Fig. 11). The pressure is read separately. The analog-to-digital converter produces a magnetic 'radiation data tape' made up of 36 words suitable for an IBM-7090 computer. In addition to the digital magnetic tape for analyses on a computer, an analog record is produced on an oscillograph for special analyses. The initial reduction of data for orbit '0' discussed later in this paper was done in this way.

The IBM-7090 computer program of the Meteorological Satellite Laboratory, U. S. Weather Bureau, requires inputs from three tapes to produce the 'final meteorological radiation tape' (Fig. 12). One source is the 'orbital data tape' containing radiation and environmental parameters in digital form. A second source is the 'orbital tape' from NASA Space Computing Center containing the position and attitude data. A third source is the calibration for converting the digital

information to meaningful physical units. The final meteorological radiation tape then is the basic repository of data from the medium-resolution scanning radiometer. To study and utilize the scanning radiometer data, appropriate computer programs must be written to 'talk' to the final meteorological radiation tape and provide for printing out data, punching cards, or producing maps. It is planned to make available to the meteorological community copies of maps showing gross results and of final meteorological radiation tapes for specialized studies. A special document will be prepared describing the contents of the final meteorological radiation tape in technical computer language for users who wish to write their own programs.

The output rate of the nonscanning, low-resolution radiometer data, because it is sampled only when TV pictures are taken, is vastly smaller than the output of the medium-resolution radiometer. Hence, the nonscanning data will be punched on cards or printed.

Preliminary results. Samples of data have been reduced by hand analyses to check the

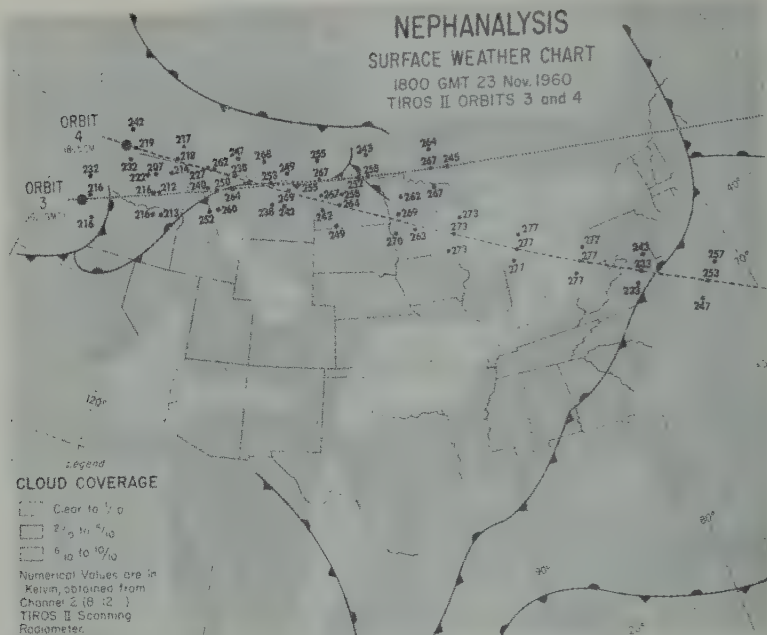


Fig. 18. Nephanalysis and frontal position map for 1800 GMT, November 23, 1960. The sub-satellite paths, the apparent blackbody temperatures viewed by the 8-12 micron channel of the medium-resolution radiometer, and the beginning times of the west-to-east passes of orbits 3 and 4 over the United States are shown. Note that maximum temperatures occur over clear areas, whereas minimum temperatures occur in the vicinity of largely overcast frontal areas (see Figs. 17 and 19).

computer program and to demonstrate the validity and usefulness of the information. Three consecutive sweeps of channel 2, recorded on the first pass between Australia and New Zealand, are shown as typical scan patterns in Figure 13. The base line corresponds to zero radiation level; at that instant both sides of the detector faced outer space. The amplitude is proportional to the energy available within the spectral range of the channel. Detectors, pre-amplifiers, and voltage-controlled oscillators contribute in the form of nonlinearities and temperature dependence to the final conversion factor between the recorded frequency deviation and the radiation level seen by the detector. Figure 14 shows all channels on a more compressed scale. Amplitude modulation on channel 7 shows sun pulses. The first five patterns are the five channels of the medium-resolution radiometer. The commutated channel 6 contains temperatures of the black and white cone and all the 'housekeeping' information. The point of verticality, where one detector sweeps a circle on

earth, can be recognized on the right, where the horizon is not intercepted at all. Figure 15 shows only the last two channels. The wider pulses on channel 7 are camera shutter signals as pictures were taken every 30 seconds. In Figure 16 the transition between the illuminated and the shadowed parts of an orbit can be identified by the cessation of sun pulses.

In the analysis, we selected samples from the first orbits over the United States and over the Tasmanian Sea. Conversion factors between frequency deviation and apparent blackbody temperatures were deduced from calibration data. The 8-12 micron radiation temperatures are plotted in Figure 17. Only the point where the optical axis intercepts the satellite path and two points 10° on either side are shown for sweeps 1 minute apart. These are less than 1 per cent of the available data points. Blackbody temperatures over the northwestern states varied from 210° to 250°K. The maximum temperature recorded in the vicinity of Ohio and California is close to 280°K. Over the eastern states tem

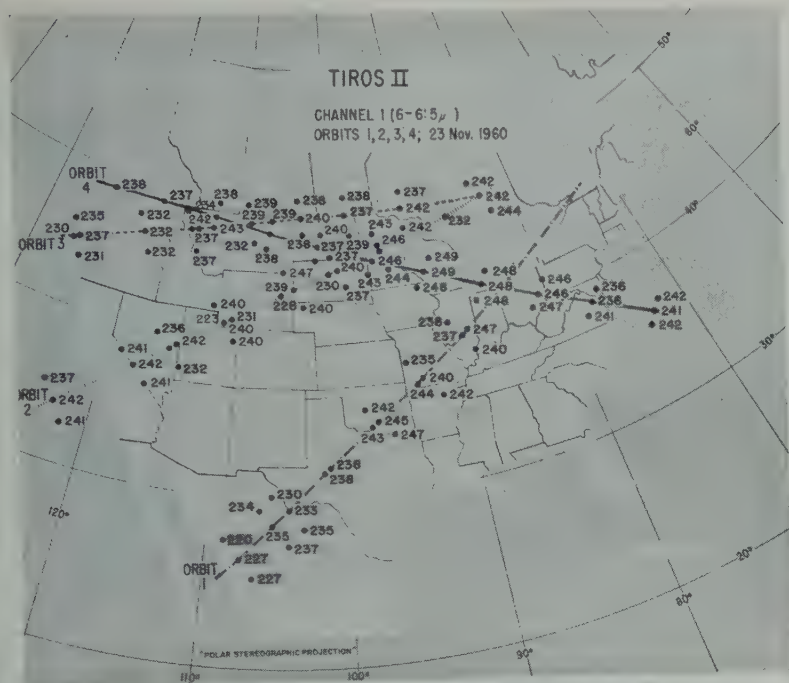


Fig. 19. Apparent blackbody temperatures viewed by the 6.0-6.5 micron channel of the medium-resolution radiometer while passing over the United States during the first four orbits after launch, November 23, 1960.

drop again. A cloud analysis and position map is shown in Figure 18. If low radiation temperatures seem to with the cloud shield very well. The micron map (Fig. 19) shows less variation in cloudy and clear areas. Radiation temperatures range from 220°K over clouds up to over cloudless areas. The absolute scale of blackbody temperatures measured by the micron channel has not yet been established with the same confidence as in the other channels. Nevertheless, the same general pattern is recognized, and results from this channel are expected to fall between those of channels 1 and 2. The relative accuracy of channels 1 and 2 is about $\pm 2^{\circ}\text{C}$, although absolute values may be up and down as much as 5°C as second-order effects in the calibration procedure are taken into account.

The area between Australia and New Zealand is reduced to the fullest extent. The Tiros position map of apparent blackbody temperatures measured around local midnight is shown in Figure 20. Low temperatures east of

the islands indicate clouds probably higher than 5 km. The maximum temperatures registered come close to the surface temperatures of water in this area.

Data from the low-resolution radiometer are shown in Figure 21, together with a nephanalysis for the first four orbits. The temperature of the black thermistor varies considerably between 306°K over Africa and 286°K over the Atlantic Ocean. The temperature of the white thermistor follows the same pattern. It shows about the same value over clear areas and values about 6° to 8°C lower over clouds. The temperature difference is smaller than expected from our calibration data. The white detector seems to act like a 'medium gray' sensor, but before final conclusions can be drawn many more data will have to be analyzed. In spite of lower temperature differences than expected, the cloud-cover analysis is in good agreement with the radiometer data.

Conclusion. The radiation experiment on Tiros II, a rather complex electronic and mechanical system, has worked very well. The instruments have produced and are continuing to

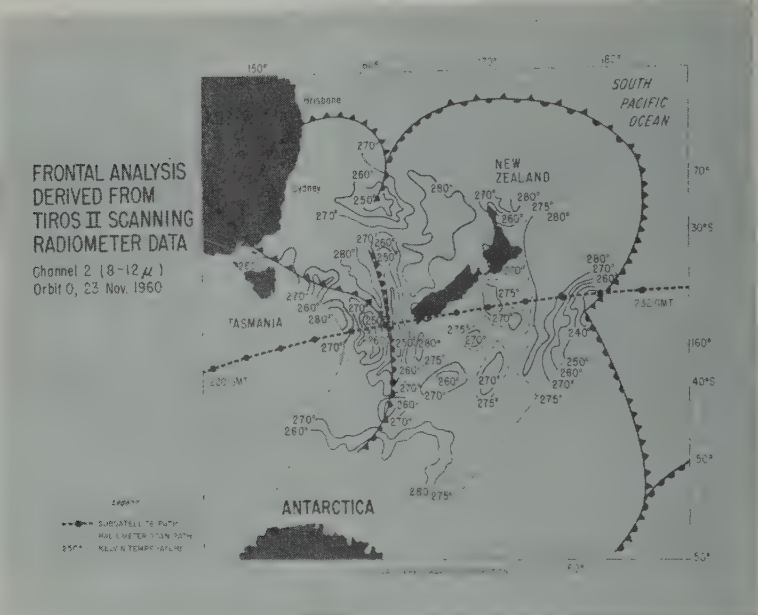


Fig. 20. Radiation map constructed from apparent blackbody temperatures viewed by 8-12 micron channel of the medium-resolution radiometer while passing over the New Zealand area during orbit '0' just after launch, November 23, 1960. The frontal positions were taken from a standard weather map, based on limited observations, and modified in accordance with the more voluminous radiation data. Market dots are placed at 1-minute intervals along the subsatellite path.

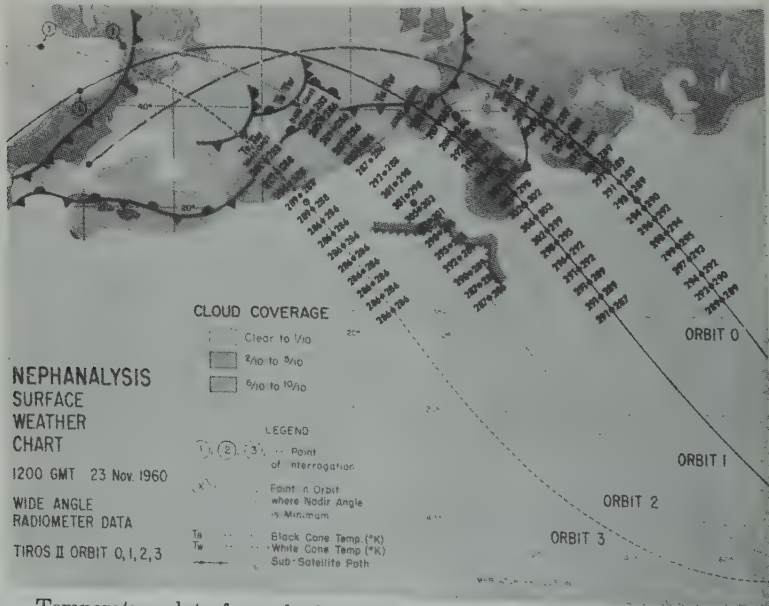


Fig. 21. Temperature data from the black and white cones of the low-resolution radiometer during orbits 0, 1, 2, and 3 over the Mediterranean, Arabia, Africa, and the Atlantic Ocean, November 23, 1960. The radiometer data are in agreement with the indicated weather data. A maximum south of Greece, whereas the two temperatures are the same over clear ocean areas where the albedo is low.

valuable data as of this writing. We make the enormous amount of data to the meteorological community as we have full confidence in all scales which the data and as soon as the automatic processing techniques are working to our own.

Acknowledgments. It is impossible to acknowledge individually the contributions of all who made this a success. We should like to thank the Engineering Company for the radiometers, the Astro-Electronics Division for the telescope system and assembling the satellite, the Optical Satellite Laboratory of the U. S. Space Bureau for its effort in the area of data processing, and our group in the Goddard Space Center who designed and built the major part of the radiation experiment instrumentation, and the calibration, and reduced the data. We should also like to thank the Optical Engineering Company of the Naval Research Laboratory for its help in the calibration and the two Command Data Acquisition Stations at Point Mugu, California, and Fort Monmouth, New Jersey, for processing the data.

REFERENCES

Greenfield, S. M., and Warren P. Manger, The motion of the Tiros I meteorological satellite due to magnetic and gravitational forces, *J. Geophys. Research*, 65(9), 2992-2995,

- Greenfield, S. M., and W. W. Kellogg, Calculations of atmospheric infrared radiation as seen from a meteorological satellite, *J. Meteorol.*, 17, 283-289, June 1960.
- Hanel, R., Low resolution radiometer, *J. Am. Rocket Soc.*, 31, 246-250, February, 1961.
- Hanel, R., J. Licht, W. Norberg, R. Stampf, and W. G. Stroud, The satellite Vanguard II: Cloud cover experiment, *IRE Trans. on Mil. Electronics*, MIL-4(2 and 3), 245-247, 1960.
- Hanel, R. A. and W. G. Stroud, Infrared imaging from satellites, *J. SMPTE*, 69(1), 25-26, 1960.
- Licht, John, and Arthur White, Polyester film belts, *Machine Design*, 32(22), 137-143, 1960.
- Sternberg, Sidney, and William G. Stroud, Tiros I: Meteorological satellite, *Astronautics*, 5(6), 32-34 and 84-86, 1960.
- Stroud, W. G., Initial results of the Tiros I meteorological satellite, *J. Geophys. Research*, 65(5), 1643-1644, 1960.
- Wark, D. Q., and G. Yamamoto, Methods of transforming terrestrial infrared radiation measurements made from satellites, unpublished paper presented at the 41st Annual Meeting of the Am. Meteorol. Soc., New York City, January 23-26, 1961.
- Wexler, R., Satellite observations of infrared radiation, *First Semiannual Tech. Summary Rept.* (Allied Research Associates, Inc., Boston, Mass.), contract AF 19 (604)-5968, AFCRC, December 24, 1959.

(Manuscript received June 15, 1961.)

A Statistical Study of Lower Atmospheric-Ionospheric Coupling

NORMAN J. MACDONALD

*Geophysics Research Directorate, Air Force Cambridge Research Laboratories
Bedford, Massachusetts*

ROBERT W. KNECHT

*CRPL, National Bureau of Standards
Boulder, Colorado*

Abstract. A study of the maximum electron density in the ionosphere, as measured by vertically incident radio wave penetration frequency, fails to show any response of this ionospheric parameter to large-scale vertical motion in the lower atmosphere.

Introduction. The possibility of coupling between the lower atmosphere and the ionized regions between 100 and 300 km has intrigued many meteorologists and ionospheric physicists for some time. Gherzi [1950] reported observation of ionospheric events preceding tropospheric air-mass changes. More recently Bauer [1957; 1958] and Mook [1958] have reported both a contemporaneous and negatively lagged ionospheric response to changes in the lower atmospheric convergence field. In a theoretical study of the upward propagation of tropospheric wave disturbances, Charney and Drazin [1961] find little evidence that lower to upper atmospheric coupling can occur, except briefly around the equinoxes.

In earlier work [Macdonald and Knecht, 1960] we had found some suggestion of an ionospheric (F₂ region) response to lower atmospheric changes. The effect appeared to occur when the convergence field and the corresponding vertical motion in the lower atmosphere (surface to 20 km) were, by our qualitative estimates, of large magnitude. The results of our study were of marginal statistical significance and showed, in part, an ionospheric effect opposite that found by Bauer and by Mook. We attributed this contradiction to differences in the nature of our experiment.

It is the purpose of this study to report an independent test of our earlier results, this time using quantitative estimates of tropospheric and stratospheric vertical motion. The study fails to confirm the earlier results.

Data and analysis. Jensen [1961] calculated

the vertical motion for several pressure levels using observed 12-hourly upper-air data, 0000 and 1200 (GCT), from 100 stations over the northern hemisphere for January and April 1958. The computations were made from observed temperature changes using the adiabatic assumption. The 12-hour time-averaged vertical velocity was computed for each specific isobaric level at each station. These data consist of the vertical motion \bar{w} , the monthly mean \bar{w} , and standard deviation σ_w at each location and were made available for purposes of this study. These data also fulfill the conditions of time independence and provide an unbiased estimate of the vertical motion. Both conditions are necessary for an experiment of this type.

The locations used for the analysis—Washington, D. C., White Sands (Albuquerque), New Mexico, and Fairbanks, Alaska—were all used in the earlier study. Anchorage, Alaska, was also studied earlier but there were no vertical-motion data computed at Anchorage so Adak, Alaska, was chosen as a substitute for this study. Churchill, Canada, was also added as another check on the earlier result, there being no reason to assume Churchill is not statistically inde-

TABLE 1

Atmospheric level	Class A	Class B
850-700 mb	—	+
300-200 mb	+	—
200-100 mb	+	—

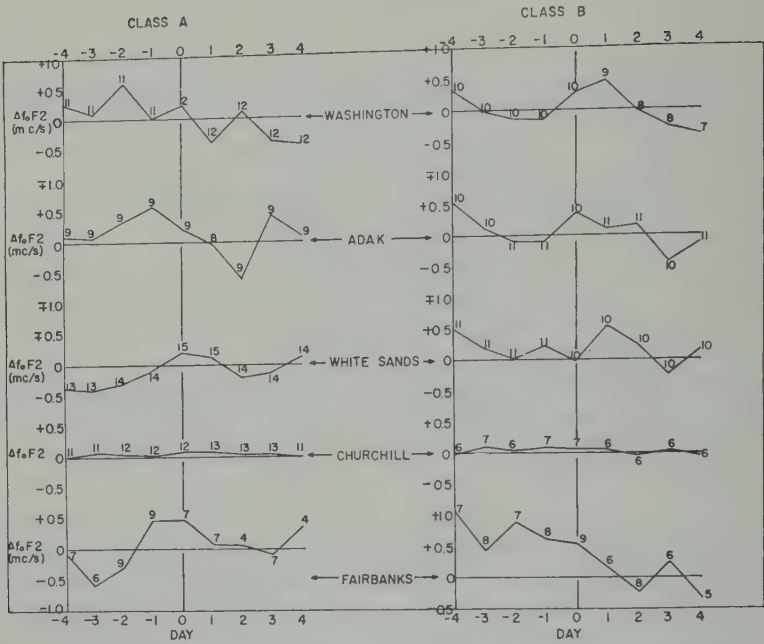


Fig. 1. The behavior of the average value of the departure from the monthly median of the penetration frequency $\Delta(f_0F_2)$ for day before (—) and after (+) the time of strong vertical motion in the lower atmosphere at each location.

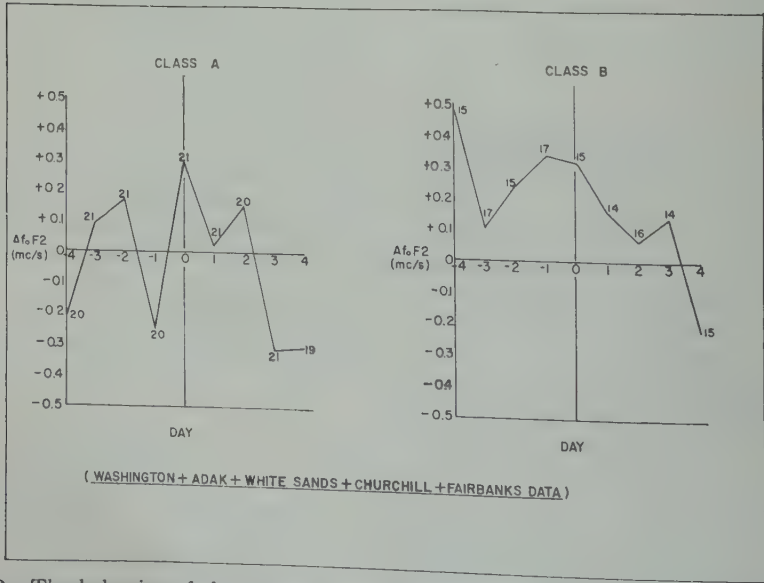


Fig. 2. The behavior of the average value of the departure from the monthly median of the penetration frequency, $\Delta(f_0F_2)$ for day before (—) and after (+) the time of strong vertical motion in at least two levels in the lower atmosphere for all locations combined.

at, in the meteorological sense, from the locations.

daily values of the F_2 -region penetration frequency, f_oF_2 , at noon (LST), 0000 (GCT), 0600 (GCT) for each month were tabulated. To circumvent the strong seasonal trend, the nature of the daily values from the monthly mean value, Δf_oF_2 , was determined in the same manner as the earlier study.

Again, the superposed epoch method of analysis was used to determine whether there were systematic changes in Δf_oF_2 around the occurrence of strong vertical motion in the troposphere and stratosphere.

In the selection of key days in this study, we considered those days when ω exceeded $\bar{\omega} + \sigma_\omega$ at one or more of the following levels:

50–700 mb ($1\frac{1}{2}$ –3 km) the lower troposphere
 300–200 mb (10–13 km) } the upper tropo-
 200–100 mb (13–16 km) } sphere and lower stratosphere

To avoid contaminating the data with ionospheric storm effects, the days of geomagnetic storm ($K_p \geq 30$) and geomagnetic storm ± 1 day were excluded from the key-day selection and the data for the superposed epoch analyses. Strong vertical motion of the same sign tends to recur for periods up to 24 hours. Therefore, to avoid the autocorrelation, further screening of key days was performed by adding the restriction that the selected dates must be separated by 24 hours or more.

Following the procedure of our earlier study, we then placed the set of key dates into two classes, depending on the distribution of the sign of the vertical motion as shown in Table 1. A positive sign indicates upward vertical motion and negative the converse.

There were no cases of $\omega \geq \bar{\omega} + \sigma_\omega$ existing in both class A and class B on the same key day.

Class A corresponds to the passage of 300-mb troughs and category A from our earlier work where we had found a suggestion of a minimum in Δf_oF_2 around the Key days. Similarly, class B in this test corresponds to condition B where we had found a suggestion of a maximum in Δf_oF_2 .

The results from the superposed epoch analyses of the quantity Δf_oF_2 listed at GCT for 24-hour intervals are shown in Figure 1. The numbers on

the graph indicate the number of cases comprising the mean for that point. The inhomogeneity of the number of cases was introduced by screening out geomagnetic storm and storm ± 1 days and the occasional lack of data from equipment problems. There is little evidence of a minimum of Δf_oF_2 near the key day in class A. In fact, a maximum of Δf_oF_2 occurs on the key ± 1 day at 4 of the 5 locations. Similarly, the maximum of Δf_oF_2 anticipated on the key day ± 1 day with class B, although found at Washington, D. C., and White Sands, is not in evidence at any of the other locations.

Another test consisted of examining the subset of key days screened for occurrences of $\omega \geq \bar{\omega} + \sigma_\omega$ at more than one level with the added condition that ω be of the proper sign at all levels. The variation of Δf_oF_2 around these key days for all locations is shown in Figure 2. Again we can find no evidence of a consistent maximum or minimum of Δf_oF_2 as our earlier work and that of Bauer had suggested.

Other superposed epoch analyses were made for Δf_oF_2 for each month separately, using local noontime values, including geomagnetic storm days, excluding marginal values of f_oF_2 and with all the various combinations, but there was still no evidence of more maxima or minima of Δf_oF_2 near the key days than one could expect by chance. We are led to conclude that the results of our earlier study are simply the results of sampling fluctuations.

Summary. Changes in the F_2 layer of the ionosphere do not show an association with vertical motion of large magnitude in the lower atmosphere. Despite these results, which negate our earlier findings and those of other workers in this field, it is possible that some coupling does exist between the F_2 region of the ionosphere and the lower atmosphere. But it must be small or nonlinear and will be difficult to demonstrate because of the large amount of 'noise' in both the meteorological and ionospheric data.

REFERENCES

- Bauer, S. J., A possible tropospheric-ionospheric relationship, *J. Geophys. Research*, **62**, 425–430, 1957.
- Bauer, S. J., Correlations between tropospheric and ionospheric parameters, *Geofis. pura e appl.*, **40**, 235–240, 1958.
- Charny, J. G., and P. G. Drazen, Propagation of planetary-scale disturbances from the lower into

- the upper atmosphere, *J. Geophys. Research*, 66, 83-110, 1961.
- Gherzi, E., Ionosphere and weather, *Nature*, 165, 38-39, 1950.
- Jensen, C. E., Energy transformation and vertical flux processes over the northern hemisphere, *J. Geophys. Research*, 66, 1145-1156, 1961.
- Macdonald, N. J., and R. W. Knecht, A possible effect of lower atmospheric divergence on the local maximum electron density in the ionosphere, *H. A. O. Inst. for Solar Terrest. Research Tech. Rept. 8*, 24 pp., 1960.
- Mook, C. P., The apparent ionospheric response to the passage of hurricane Diane (1955) & Washington, D. C., *J. Geophys. Research*, 63, 569-570, 1958.

(Manuscript received August 8, 1961.)

Upper-Atmosphere Structure Measurement Made with the Pitot-Static Tube

J. E. AINSWORTH,¹ D. F. FOX, AND H. E. LA GOW¹

*Goddard Space Flight Center
Greenbelt, Maryland*

Abstract. Profiles of atmospheric pressure, density, and temperature for the region from 20 to 100 km above Fort Churchill, Canada, were computed from the data obtained during a single 31-day rocket flight. The temperature profile had a major maximum at 58 km, a secondary maximum at 82 km, and several minor maxima in the region from 20 to 55 km. During the same flight a horizontal-wind profile was obtained in the region from 80 to 115 km. Average wind speed was 15 m/sec, and the maximum wind speed was 250 m/sec at 115 km. The method of measurement, the problems encountered, and the results are discussed. Methods are prescribed for the use of the pitot-static tube for accurate synoptic atmospheric structure measurements from 20 to 80 km and for the measurement of semidiurnal pressure variations in the same region.

INTRODUCTION

The first extensive measurements of upper-atmosphere pressure, density, and temperature were made possible immediately after World War II by the availability of high-altitude rockets for research. As might be suspected, the atmospheric structure information was obtained from these rockets by aerodynamic methods and was based on the measurement of air pressure at various positions on the rocket's surface. In the particular method used by Koll, and LaGow [1952], atmospheric density was computed from a measurement of impact pressure at the nose of the rocket. Atmospheric pressure was computed from a pressure measurement at that position along the body of the rocket where the surface pressure was roughly equal to the ambient atmospheric pressure. In this method the entire rocket body was used as though it were a pitot-static tube. At high altitudes, where the mean free paths of the component gases of the atmosphere were greater than the rocket diameter and where the moving rocket was moving sidewise, the presence of a gage opening located on the side of the rocket experienced a 'spin modulation.' This method accurately measured the ambient atmospheric pressure twice during each rocket spin period. The ambient atmospheric density was

computed from the maximum pressure excursion measured during each spin period.

Subsequent to the above measurements, LaGow [LaGow and Ainsworth, 1956] made pressure, density, and temperature measurements in the Arctic and demonstrated that aerodynamic methods similar to those above could be used with small rockets.

It was apparent from this early rocket work that these particular aerodynamic methods had considerable promise for future measurements: (1) The measurements could be made continuously with altitude so as not to omit any detail in the structure of the atmosphere; (2) the measurements could be made from the ground to the beginning of the *F* region, and thus it was possible to determine the vertical extent of and to study the relations between the various phenomena observed over this large altitude range; (3) they could be made by day or night and practically irrespective of weather conditions and would thus be able to present representative structure information; (4) certain redundancy of data was obtained and was useful for establishing the consistency and accuracy of the measurements at any time that unusual effects were observed; and (5) the methods were adaptable to small rockets. It was clear, also, that a horizontal wind profile might be extracted from the measured pressures provided that reasonable improvements in instrumentation were accomplished.

It was also apparent that a number of problems

¹Formerly with the U. S. Naval Research Laboratory, where the work reported herein was done.

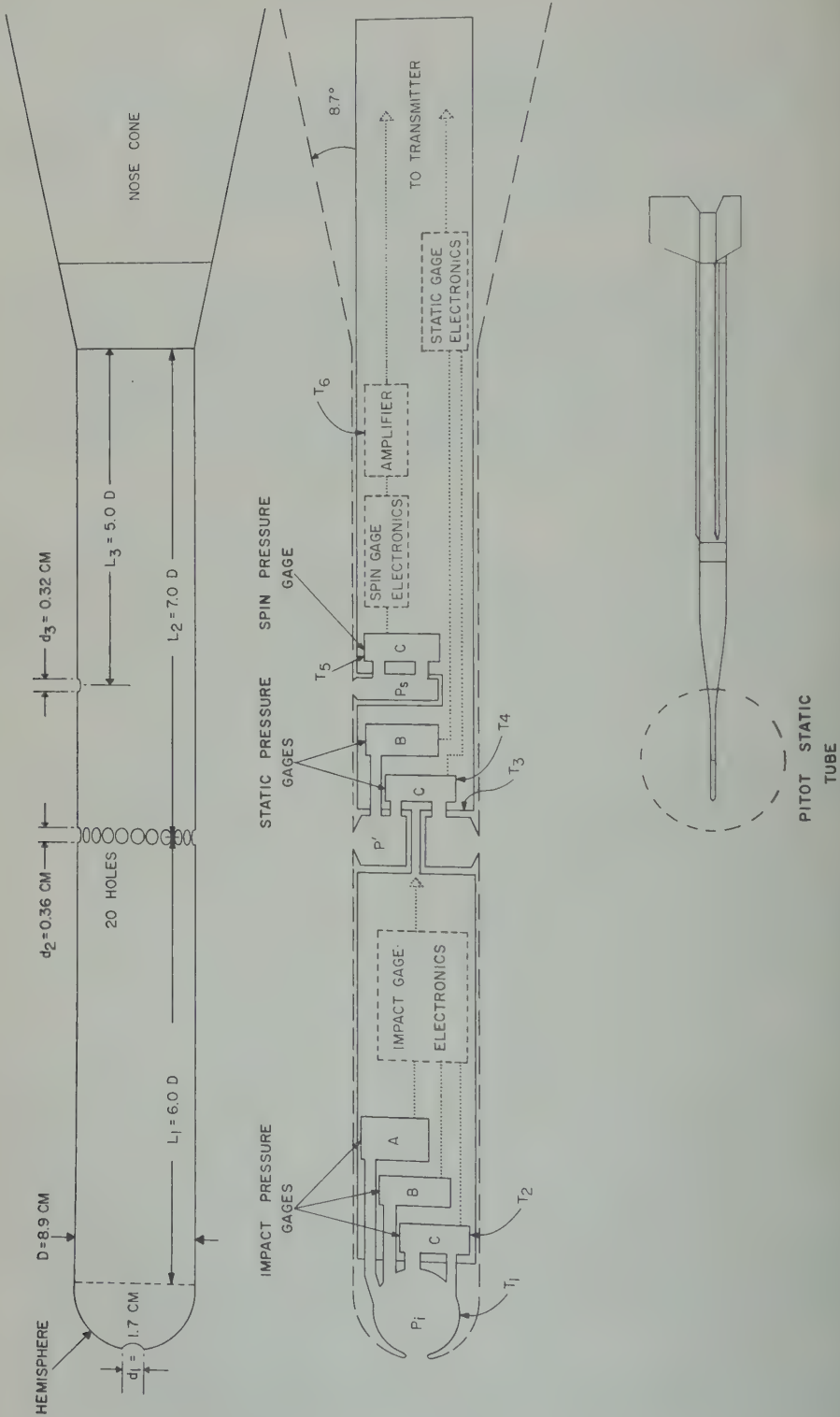


Fig. 1. Mechanical and electrical configuration of the pitot-static tube.

LE 1. Rocket Flight Data for Aerobee-Hi NN3.15F Launched October 31, 1958, at 1359 CST

Altitude, km	Ascent Time, sec	Descent Time, sec	Rocket Speed, km/sec	Knudsen Number K_D	Reynolds Number Re_D	Mean Free Path λ , cm
0	39.0		1.05	1.04×10^{-5}	5.69×10^5	9.31×10^{-5}
5	43.2		1.27	2.38	2.94	2.13×10^{-4}
10	46.9		1.52	5.14	1.63	4.59
15	50.0		1.71	1.12×10^{-4}	8.33×10^4	9.98
20	53.0		1.69	2.36	3.90	2.11×10^{-3}
25	56.0		1.67	5.14	1.72	4.59
30	59.1	403.3	1.64	1.12×10^{-3}	7.27×10^3	9.98
35	62.0	400.2	1.61	2.24	3.46	2.00×10^{-2}
40	65.4	397.0	1.58	4.24	1.77	3.79
45	68.7	393.8	1.55	7.75	1.01	6.92
50	72.0	390.4	1.52	1.46×10^{-2}	5.60×10^2	1.30×10^{-1}
55	75.4	387.0	1.48	2.98	2.71	2.66
60	78.8	383.6	1.45	6.45	1.29	5.76
65	82.4	380.0	1.42	1.29×10^{-1}	5.88×10^1	1.15×10^0
70	86.0	376.3	1.38	2.62	2.99	2.33
75	89.8	372.5	1.35	5.79		5.15
80	93.6	368.7	1.31	1.24×10^0		1.55×10^1
85	97.6	364.7	1.28	2.77		2.46
90	101.7	360.6	1.24	5.76		5.12

ed to be solved in future applications of method. To obtain the maximum possible information about the structure of the atmosphere the subsequent measurements would be required to have greater accuracy, greater altitude, and a greater altitude range than the exploratory measurements. In addition, a number of flights would be required if the dependence of pressure, density, temperature, and winds upon latitude, season, time of day, and solar activity was to be obtained. For large numbers of flights there is the problem of reduction of cost by means of simplifications of equipment, and analysis. However, simplifications must not be allowed to reduce the range and accuracy of the measurements.

The work reported here is an extension of the Aerobee-Hi rocket experiments. We shall describe the methods employed, the problems encountered, and the results obtained from the flights of Aerobee-Hi rockets equipped with pitot-static tubes and launched from Fort Churchill, Manitoba, during the period from 1956 to 1958. Since the total number of flights was small, and atmospheric structure information was limited and a number of the problems associated with second-generation measurements were solved and clarified. The main results of the three flights

are as follows: (1) Pressure measurements have been used for the first time to obtain an upper atmosphere wind profile. (2) A temperature profile with improved detail has been obtained from 20 to 110 km. (3) A certain regularity in altitude spacing has been observed for the maxima of the wind and temperature profiles.² (4) It is clear from the results that accurate machine-computed density and temperature profiles can be obtained in the region from 20 to 80 km by the measurement of impact pressure alone, and by means of simple available pressure sensors. (5) It is clear from the results that accurate measurements of semidiurnal pressure changes can now be made by means of a static-pressure manifold and simple available pressure sensors.

Unless otherwise stated, the paper will describe the last of the three flights. This flight was the most successful, and the most recently developed instruments were used.

INSTRUMENTATION

Gages. The mechanical and electrical configuration of the pitot-static tube and its gages is shown in Figure 1. Gage A was a commercial capsule-potentiometer unit with a full-scale

² To be discussed in detail in a subsequent paper.

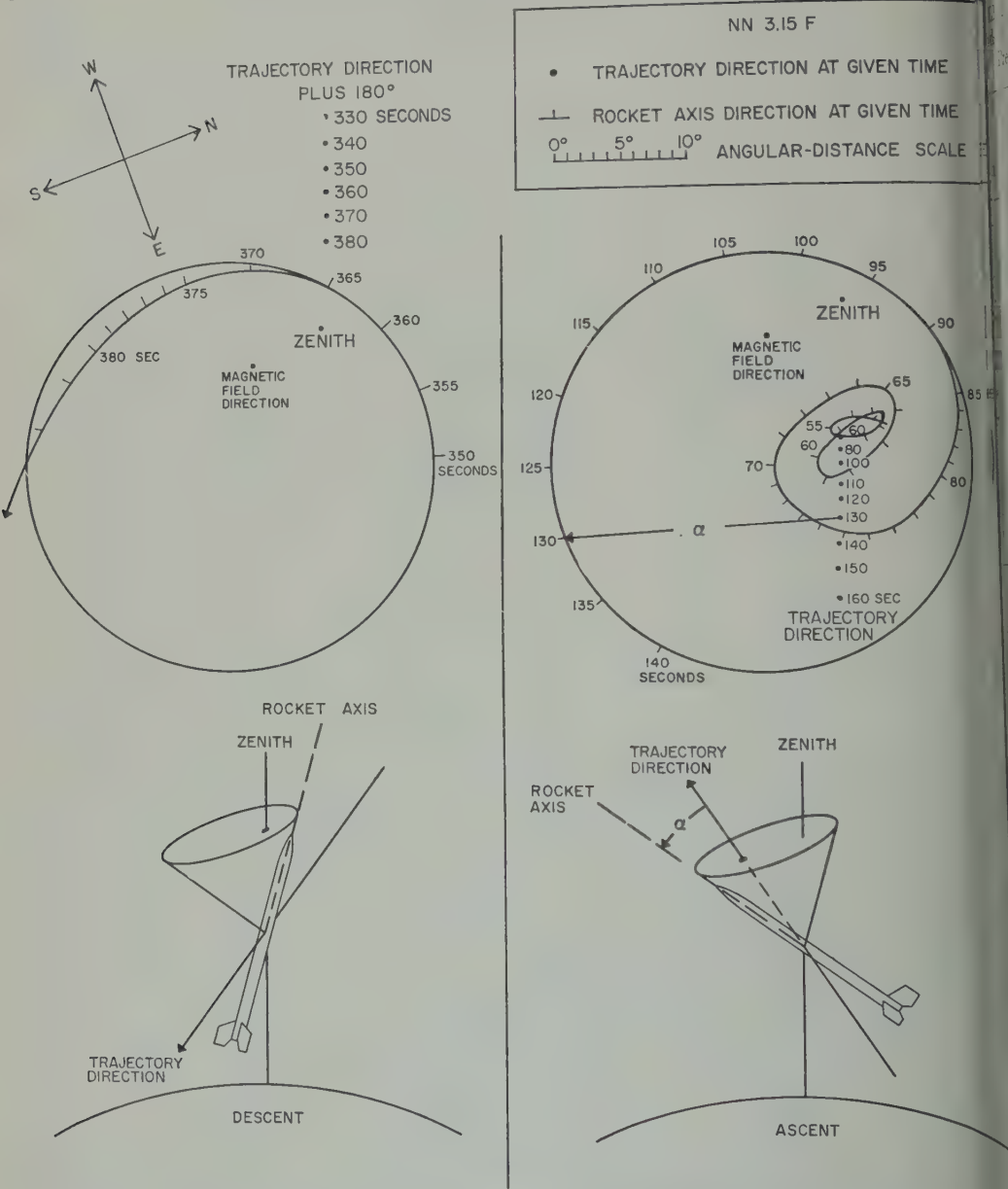


Fig. 2. Precession cone for NN3.15F. At a given time, the rocket angle of attack α is obtained by joining the trajectory and precession cone points for that time. An example is shown in the top right polar diagram at a flight time of 130 seconds.

pressure of 760 mm Hg. Gages B and C were specially developed for pitot-static tube use [Decker Aviation Corp., 1957] because there were no gages available with sufficient range, accuracy, stability, ruggedness, and insensitivity to over-pressure, and with sufficiently fast pres-

sure response. Gages B and C incorporated diaphragm pressure sensors with ranges of 20 mm Hg and 0.2 mm Hg, respectively. They were used to measure pressures from 20 mm Hg to $5(10)^{-4}$ mm Hg. Gage C was also used with an amplifier to measure spin pressure modulation

2. The Effect of Horizontal Atmospheric on the Density Obtained from the Impact Pressure Measurement for NN3.15F

Fictitious Wind		Correction to ρ , %
$ W $, m/sec	Azimuth, deg	
37	125	+ .9
50	110	+1.0
61	105	+1.1
70	80	+1.1
120	120	+2.3
110	190	+0.6
110	125	-2.8
Measured values		-2.4
of the horizontal		+0.6
winds are used		+1.4
for computing the		-0.6
correction for		-0.3
80 to 110 km.		-0.4
See text.		-1.2

an amplitude as small as 10^{-6} mm Hg. All gages used mechanical pressure sensors, and measurements of true pressure were obtained irrespective of the composition of the gas within the gage. Gages B and C will not be described since they have been recently supplanted by gages with improved characteristics [Decker, 1961]. Because of the wide range of pressures to be measured and the accuracies required, particular attention was paid to field calibration of gages B and C, and special equipment and methods were devised for [Lanick and Ainsworth, 1961a, b; Ainsworth, and LaGow, 1961]. Pitot-static tube shell temperatures were measured at points T_1 to T_6 in Figure 1.

System response times. In conventional applications of the pitot-static tube, long pipes are used to connect the impact and static pressure sensors with their respective gages located in the body of the vehicle. In rocket use of the pitot-static tube the impact and static pressures are so rapidly that long pipes cause pronounced pressure lags at the gages. The conductance of long pipes is particularly poor at high altitudes where conditions of free molecular flow exist. Because of this fact, gage outgassing may be completely the high-altitude pressure measurements. For the present experiments particular difficulties were avoided by locating the gages inside the pitot-static tube

shell and next to their respective chambers as shown in Figure 1. The impact and static pressures changed by as much as 25 per cent per second, but due to the proximity of the gages to their chambers and due to suitable gage design the pressure lag at the gages was less than 0.5 per cent. The pressure lag in following 'spin modulated' pressures was also less than 0.5 per cent [Schaaf and Cyr, 1949; Ainsworth and LaGow, 1956].

Gas interference. Particular care was given to the reduction of gas interference. Gage and chamber pressure-response time constants were made as small as possible. The interior surfaces of the gages, the surfaces of the chambers, and the exterior surface of the pitot-static tube were highly polished and repeatedly outgassed. All pitot-static tube welds and seals were checked by means of a helium leak detector, the entire rocket body was sealed, and the rocket engine propellant cutoff valves were closed after burn-out.

No evidence of residual gas was seen in the measurements, although a residual gas pressure of 10^{-4} mm Hg at 110 km would have been readily detected. Ion gage measurements on the same flight indicated that residual pressures under the above conditions were likely to be $2(10)^{-6}$ mm Hg or less.³

ROCKET TRAJECTORY AND ASPECT

Rocket trajectory was obtained by means of Dovap [Newell, 1953] trajectory information, and the trajectory error was less than 0.05 km on the ascent and 0.1 km on the descent. The basic flight information is presented in Table 1.

Rocket aspect was obtained by means of an optical aspect system [Kupperian and Kreplin, 1957] and a single magnetometer. Canting of the rocket fins gave the rocket a final spin of 1.71 rps. The rocket's precession cone (Fig. 2) had a vertex angle of $32.8 \pm 0.1^\circ$ and the precession period was 89.6 ± 0.1 sec. The maximum errors in the location of the cone's axis and in the subsequent location of its time scale reference were both less than 1° .

IMPACT PRESSURE

The impact pressure measurements obtained in flight are shown in Figure 3.

³ This information was furnished by R. Horowitz, who conducted the high-altitude experiment.

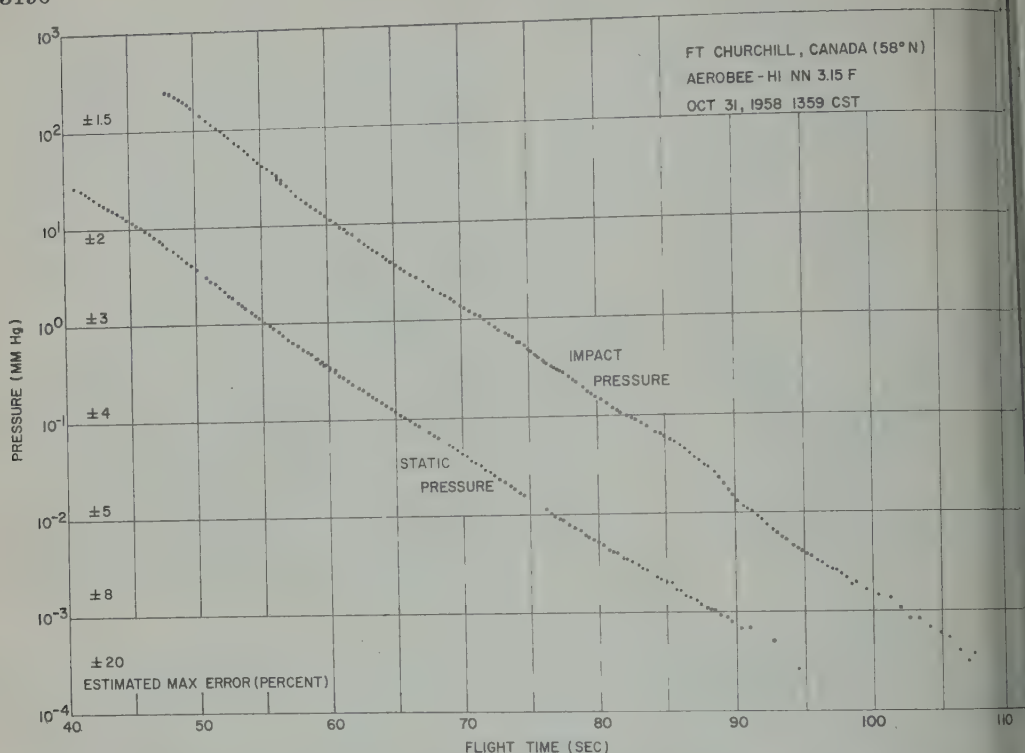


Fig. 3. Impact and static pressure data before aerodynamic corrections were applied.

At low altitudes, impact pressure was converted to density by means of the Rayleigh pitot formula for a diatomic gas [Newell, 1953]:

$$\rho = (0.144P_i - 0.066P)/|\mathbf{V}_T|^2, \quad (1)$$

$$[g/m^3 = mm \text{ Hg}/(km \text{ sec}^{-1})^2]$$

where P_i is the ideal impact pressure, P is the ambient pressure, and \mathbf{V}_T is the wind created by the motion of the rocket along its trajectory.

A number of corrections to the measured impact pressure were required in order to obtain the ideal impact pressure. The first correction was required because of the high gas temperatures which exist in the vicinity of the nose of the pitot-static tube at high Mach numbers. These high temperatures cause a reduction in the ratio of the specific heats which is known as caloric imperfection [Hill, Baron, and Schindel, 1956] and which causes the measured impact pressure to be greater than the ideal impact pressure. For our particular Mach numbers, the values of the measured impact pressure were determined [Ames Research Staff, 1953] to be greater than the ideal impact pressure by from 0.25 to 0.65

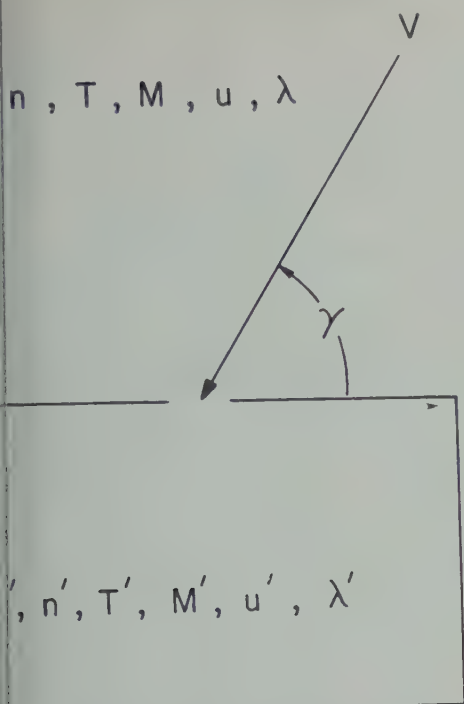
per cent in the region from 20 to 30 km and by approximately 0.8 per cent in the region from 32 to 80 km.

A second correction to the measured impact pressure was a reduction in value of 0.2 to 0.3 per cent to correct for pressure lag within the gage and its chamber.

A third correction was for impact chamber orifice size. According to the work of Chamberlain and Schaaf [1954], the size of the orifice of this particular impact chamber caused the measured impact pressure to be lower than the ideal impact pressure by less than 0.3 per cent in the region of continuum flow.

A fourth correction to the measured impact pressure was for viscous effects. Starting at 7 km and at a Reynolds number with respect to diameter (Re_D) of 700, correction was made according to the data obtained by Sherman [1953] for a 5 to 1 source-shaped tube. The correction was about 1 per cent at 82 km and 9 per cent at 89 km where an abrupt transition to free molecular flow began.

A fifth correction to the measured impact



4. Environment and pressure chamber notation.

There was for rocket angle of attack α and based on wind-tunnel tests of the angle of sensitivity of the pitot-static tube [Laur-1958]. With this correction (1) becomes

$$(0.144P_i/\cos^{1/2}\alpha) - 0.066P/|V_T|^2 \quad (2)$$

The cosine term presents an approximate correction which is accurate to about ± 0.5 per cent. $\alpha = 13^\circ$. Maximum correction to P_i was ± 0.7 per cent at 90 km. At 80 km and below, the correction was less than ± 0.2 per cent. Atmospheric winds must also be accounted for in the use of (2). Atmospheric wind vector W must be added to rocket-wind vector V_T to find the total wind vector V which replaces V_T in (2). The angle β between V and the rocket's longitudinal axis is the effective angle of attack and must be used in (2) to replace α , the actual angle of attack. The calculation of the effect of atmospheric winds upon the corrected density is simplified if W is separated into vertical and horizontal components, and the effect on (2) of each component considered separately.

The first-order correction to (2) for the vertical

atmospheric wind component is made by replacing the denominator by $[|V_T| - |W_v| \cos \psi]^2$, where ψ is the angle between the vertical wind vector and the rocket's longitudinal axis. From Table 1 and Figure 2 it can be seen that a vertical wind of 7 m/sec would cause a negligible change in the rocket's angle of attack but would require a 1 per cent correction to the denominator of (2).

We have found no reference to vertical wind speed measurements in the region from 20 to 80 km. According to Murgatroyd [1957] the wind in the region from 80 to 100 km 'is substantially horizontal but individual observations may give up to 10 m/sec vertical components.' Root-mean-square values of the vertical wind in the region 80 to 100 km have an upper limit [Greenhow and Neufeld, 1959] of 1 to 2 m/sec at Jodrell Bank ($53^\circ N$). In this work we have assumed that these vertical wind speeds are also representative of the entire region from 20 to 90 km. With this assumption the maximum error in the density measurements due to vertical winds could be as much as 1.5 per cent, but the root-mean-square error would be 0.15 to 0.3 per cent.

The solution for the effect on the measured density of a horizontal wind component is straightforward but is not included here because of its length. It was not possible to measure a horizontal wind profile from 20 to 78 km by means of the instruments used on the flights described here. To estimate the maximum possible size of corrections to the measured density, corrections were computed (see Table 2) for a fictitious horizontal wind. This fictitious wind comprised the maximum winds measured from 20 to 60 km by the autumn grenade flights SM1.01, SM1.07, and SM1.08 at Fort Churchill, Canada [Bandein, Griffith, Nordberg, and Stroud, 1959], and was joined at 78 km to the horizontal wind measured by means of the pitot-static tube. Above 78 km actual measurements of the horizontal wind (see Fig. 6) were used to compute the corrections to the density.

To reduce the effect of the horizontal wind on the impact pressure measurement, the zenith angle of the trajectory at the beginning of flight should be kept small. The correction of -2.8 per cent at 75 km (see Table 2) is associated with an initial zenith angle of about 9° . The correction would be reduced to -1.7 per cent for an initial zenith angle of 4.5° .

The absence of significant low-altitude vertical and horizontal winds is suggested by the excellent agreement obtained during the first and second flights between the density profiles computed from the pitot-static tube measurements and the density profiles obtained from their related radiosonde measurements. On both of these flights impact pressure measurements, radiosonde temperature measurements, and radiosonde hypsometer pressure measurements were obtained from 19 to 28 km. The density curves obtained from the rocket flights and the related density curves obtained from the radiosonde measurements had average separations of less than 1 per cent, and there were no systematic differences between related curves.

The transition of the measurement of impact pressure to conditions of free molecular flow is shown in Figure 3 and was characterized by a reduction of impact pressure to approximately one-half of the extrapolated value for conditions of continuum flow. It was judged that at 100 km the measured impact pressure was within 5 per cent of the downward extrapolation of the free molecular flow measurements of impact pressure.

Conversion from free molecular flow impact pressure to ambient density was obtained from the following considerations. Figure 4 depicts a region containing a gas characterized by mass density ρ , number density n , temperature T , mean molecular mass M , a most probable thermal velocity u , mean free path λ , and a mass velocity \mathbf{V} relative to the region inside a cavity where the several properties are designated by corresponding primed symbols. The direction of the mass velocity intersects the orifice surface of the cavity at angle γ . The orifice has area A and a diameter $d \ll \lambda$. Outside dimensions of the cavity and the gas density within the cavity are such that a gas molecule approaching the orifice experiences negligible interference from gas molecules given off from the cavity's exterior surface or emerging from the orifice. Under these conditions the number rate of flow of gas molecules into the cavity is given by [Sanger, 1950; Tsien, 1946; Wiener, 1949; Schultz, Spencer, and Reifman, 1948]

$$N = (Anu/2\sqrt{\pi})\{\exp(-S^2) + S\sqrt{\pi}[1 + \operatorname{erf}(S)]\} \quad (3)$$

$$= (Anu/2\sqrt{\pi})F(S) \quad (4)$$

where

$$\operatorname{erf}(S) = \frac{2}{\sqrt{\pi}} \int_0^S e^{-x^2} dx$$

$$S = (|\mathbf{V}| \sin \gamma)/u \\ = (|\mathbf{V}| \sin \gamma)/(2RTM^{-1})^{1/2}$$

and the molar gas constant is

$$R = 8.315 \times 10^7 \text{ (ergs)(deg}^{-1} \text{ K)(g-mole}^{-1} \text{)}$$

S is taken as positive when \mathbf{V} is directed into the orifice and negative when \mathbf{V} is directed away from the orifice.

Inside the cavity are gas molecules having the outside region as their source, and in addition there are molecules evolved during outgassing of the interior walls of the cavity. When the mean free path λ' of gas in the cavity is greater than the interior dimensions of the cavity and when the orifice diameter is small compared with the interior dimensions of the cavity, so that gas entering from the outside region has adequate time to acquire the temperature T' of the gas on the wall, then (4) may be used to compute the return rate of flow N' of outside-region gas molecules from the cavity. Since $\mathbf{V} = 0$, $F(S) = 1$ and (4) becomes

$$N' = An'u'/2\sqrt{\pi} \quad (7)$$

Under equilibrium conditions the rates of outside-region mass flow into the chamber and out of the chamber are equal and require that

$$NM = N'M' \quad (8)$$

Equations 4 and 7 and the equation of state are used in (8) to obtain the pressure ratio

$$P'/P = (M/M')^{1/2}(T'/T)^{1/2}F(S) \quad (9)$$

Should equilibrium be sharply disturbed by a large wind gradient or by a rapid change in mean molecular mass, the time required to reach a new equilibrium depends upon the time constant of the impact pressure system. In this instance the free molecular flow time constant is about 0.01 seconds, and 99.99 per cent of any required step change is accomplished in about 10 time constants or 0.1 seconds (about 10 meters in altitude).

In the absence of appreciable outgassing, (9) may be used to compute ambient density from the measured impact pressure P , by taking

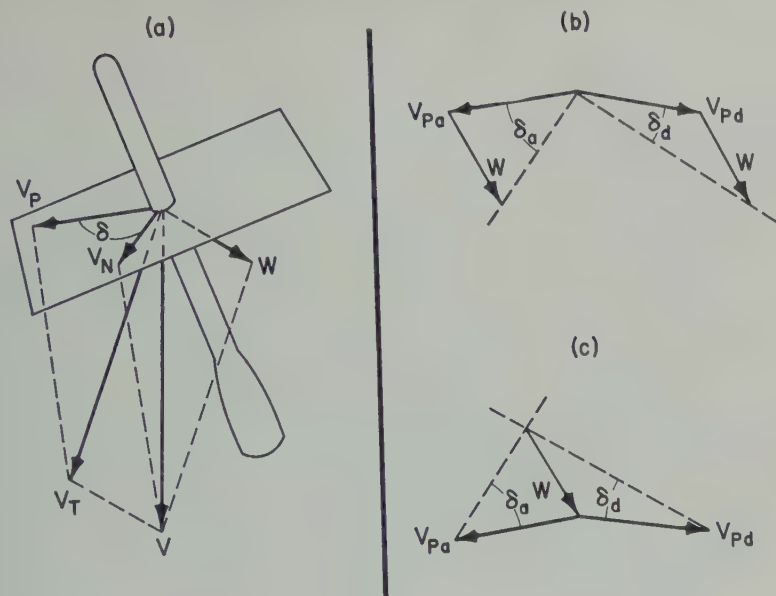


Fig. 5. Wind computation geometry (see text).

ference

$$-P'(-S) = P(M/M')^{1/2}(T'/T)^{1/2} \cdot [F(S) - F(-S)] \quad (10)$$

the high speeds obtained in flight, S was small. As a consequence, the term $P'(-S)$ was less than 0.05 per cent of $P'(S)$ and was thus neglected. Equation 10 and the equation of state yield

$$P'(S) = P_i \approx \sqrt{\pi} \rho u' S \quad (11)$$

$$\rho \approx P_i / (\sqrt{\pi} u' |V| \sin \gamma) \quad (12)$$

$$\approx P_i / (\sqrt{\pi} u' |V| \cos \beta) \quad (13)$$

where β , the complement of γ , is the effective impact angle of attack as viewed from the orifice of the impact-pressure chamber.

Because of the vector component form of the denominator of (13), the effect of atmospheric winds can be readily taken into account. Equation 13 can be written

$$P_i [(\sqrt{\pi} u') (|V_T| \cos \alpha - |W_v| \cos \psi - |W_H| \sin \theta \cos \phi)]^{-1} \quad (14)$$

where the three terms of the denominator are, respectively, the components along the rocket's longitudinal axis of the rocket wind, the vertical atmospheric wind, and the horizontal atmos-

pheric wind. The angle α is the rocket angle of attack, ψ is the angle between the vertical wind vector and the rocket's longitudinal axis, θ is the angle of the trajectory from the zenith, and ϕ is the angle between the horizontal wind vector and the rocket's azimuthal direction.

It can be seen from (14), Table 1, and Figure 2 that a vertical wind of 12 m/sec would cause a 1 per cent change in ρ . Since the maximum vertical wind cited by Murgatroyd was 10 m/sec the effect of vertical winds on the computed ρ is small. Horizontal winds were measured (see Fig. 6) and the corrections to ρ shown in Table 2 were made.

STATIC PRESSURE

The static pressure measurements are shown in Figure 3.

Before the first flight there was concern about the possibility that at low densities the shock wave at the flare where the pitot-static tube joined the nose cone would separate the boundary layer on the tube and would move forward along the tube to create anomalous pressures, first at the spin-pressure chamber and then at the static-pressure chamber. Subsequent wind-tunnel tests [Laurmann, 1958] at $5000 > Re_D > 90$ indicated no shock wave separation or anomalous pressure with the flare as close as two diameters from

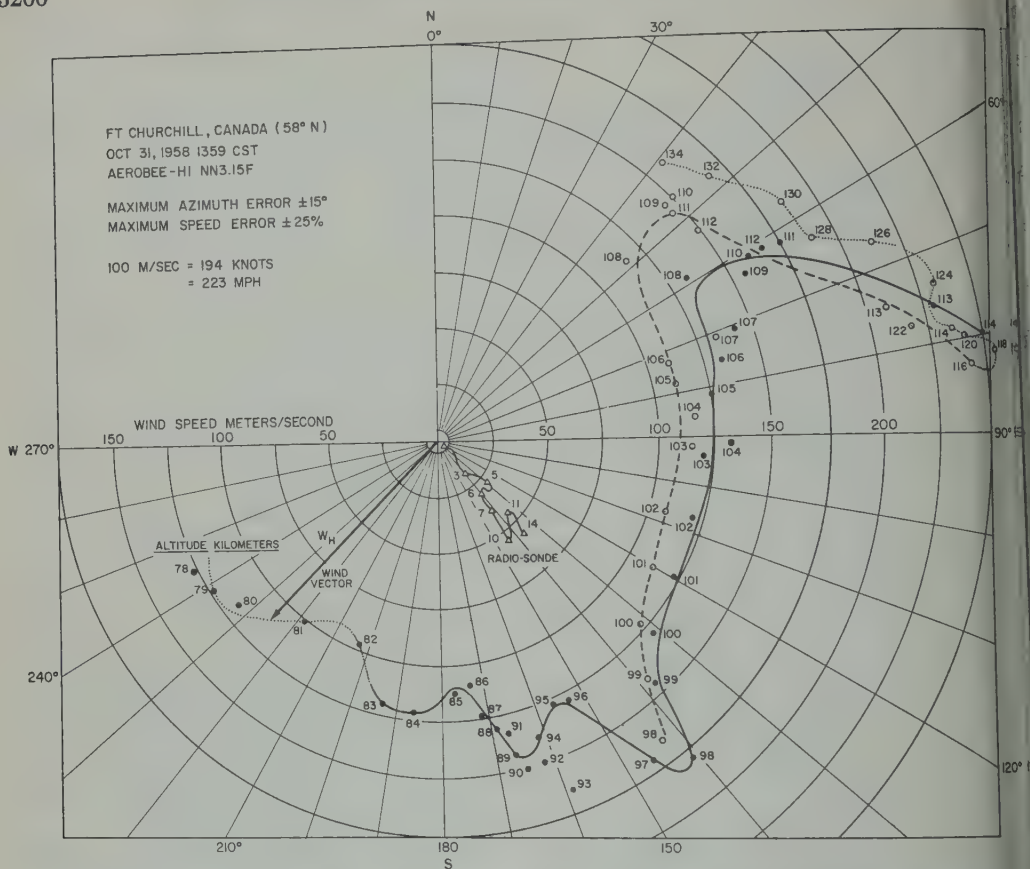


Fig. 6. Horizontal-wind profile above Fort Churchill. The solid curve was obtained from ascent and descent spin pressures. The dashed curve was obtained from the ascent spin pressure and impact pressure. The dotted lines are extrapolations.

the static-pressure chamber orifices. In flight no anomalous effects were observed at either pressure chamber.

In general, the value obtained for the static pressure is lower than ambient pressure by an amount depending on the ratio L_1/D (see Fig. 1) and on the Mach number. It was not possible to determine this static pressure deficiency $\Delta P'$ of our specific pitot-static tube with wind-tunnel tests. It was possible during flight, however, to compare the static pressure over the interval from 18 to 31 km with the ambient pressure measured by the radiosonde over the same interval. The results gave $100 (\Delta P'/P) \approx 4.3$ per cent independent of Mach number for $3.75 < M < 5.6$. These results were obtained on the first flight, during which the hypsometer radiosonde pressures were quite accurate and the

static pressure measurements may have been several per cent in error. On the third flight, when static pressures were measured with better accuracy, the radiosonde pressure measurements were poor and a direct comparison of pressures was unsatisfactory. The method (described later) for determining atmospheric temperature on the third flight suggested a value of $100 (\Delta P'/P) = 2.0 \pm 2.0$ per cent which was independent of Mach number. This 2 per cent correction was made on all static-pressure data up to about 85 km, at which height it appeared that continuum flow no longer existed.

Wind-tunnel-determined static pressure corrections for rocket angle of attack were available [Laurmann, 1958] for the region from 20 to 85 km. The corrections at 20 and 85 km were 0.1 and 0.8 per cent, respectively. The effect on the

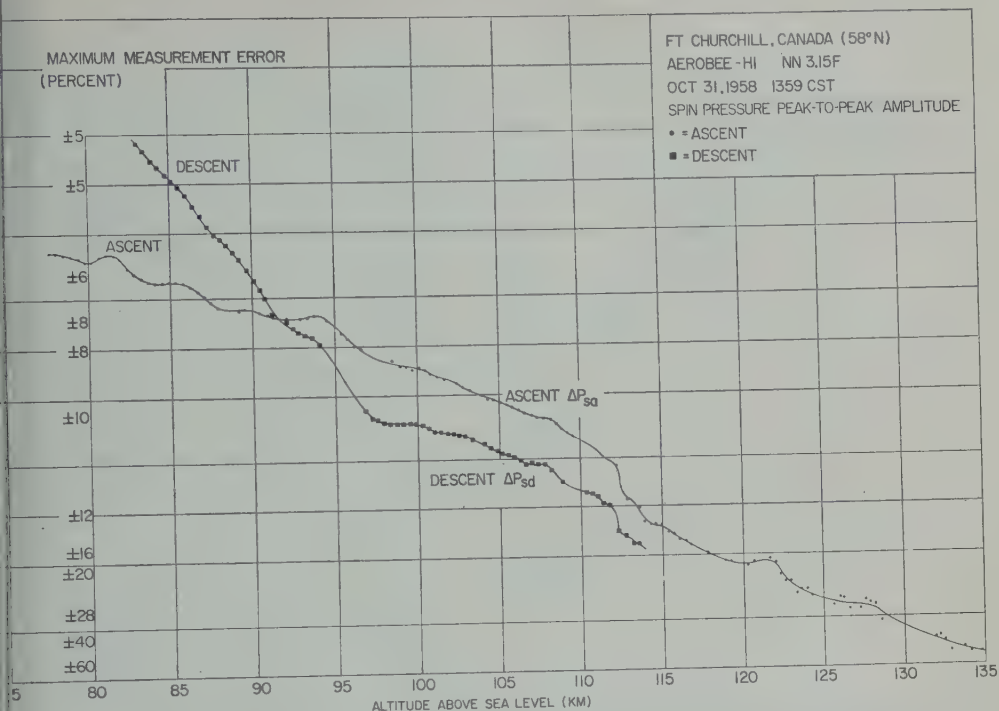


Fig. 7. Peak-to-peak amplitude of the spin pressure signal. Different angles of attack of the rocket create peak-to-peak spin pressure amplitudes that differ on ascent and descent. Equal wind shear at the same altitudes on ascent and descent create matching slope changes in the modulation amplitudes.

pressure of the maximum possible vertical was shown to be negligible. The correction for the fictitious horizontal wind from 20 to 80 km (see preceding section) was also negligible. The correction for the measured horizontal wind from 80 to 90 km opposed the correction for the angle of attack and resulted in a net correction of less than 0.3 per cent to the static pressure.

In the region of free molecular flow the correction of P' to ambient pressure P was based on the integration of (3). In the integration it was assumed that the chamber orifice was a thin slit passing the chamber, and the total mass flow out of the chamber was equated with the total mass flow out of the chamber. This initial simplifying assumption of a slit orifice is believed to be valid since the area of the slit and the area of the actual static-chamber orifices were both small in comparison with the total interior area of the chamber and there was a negligible probability of a molecule entering and then leaving the chamber before it had assumed the

chamber temperature. Since the effective angles of attack were small, further simplification of the integration was possible by use of only the first few terms of the power series expansion of (3). In the resultant equation,

$$P \approx P'(M'/M)^{1/2}(T/T')^{1/2} \cdot [1 + (1/2)(|V| \beta/u)^2]^{-1} \quad (15)$$

the approximation to (3) resulted in an error of less than 0.5 per cent in P at $\beta = 7^\circ$. The effective angles of attack were determined during the computation of the horizontal atmospheric wind as described in the next section.

Vertical winds were shown to have a negligible effect on the effective angle of attack.

SPIN PRESSURE

Winds. The spin pressure data used for determining winds were (1) the phase of the spin pressure maximum, and (2) ΔP_s , the peak-to-peak amplitude of the spin pressure. In the absence of an atmospheric wind, the phase of a

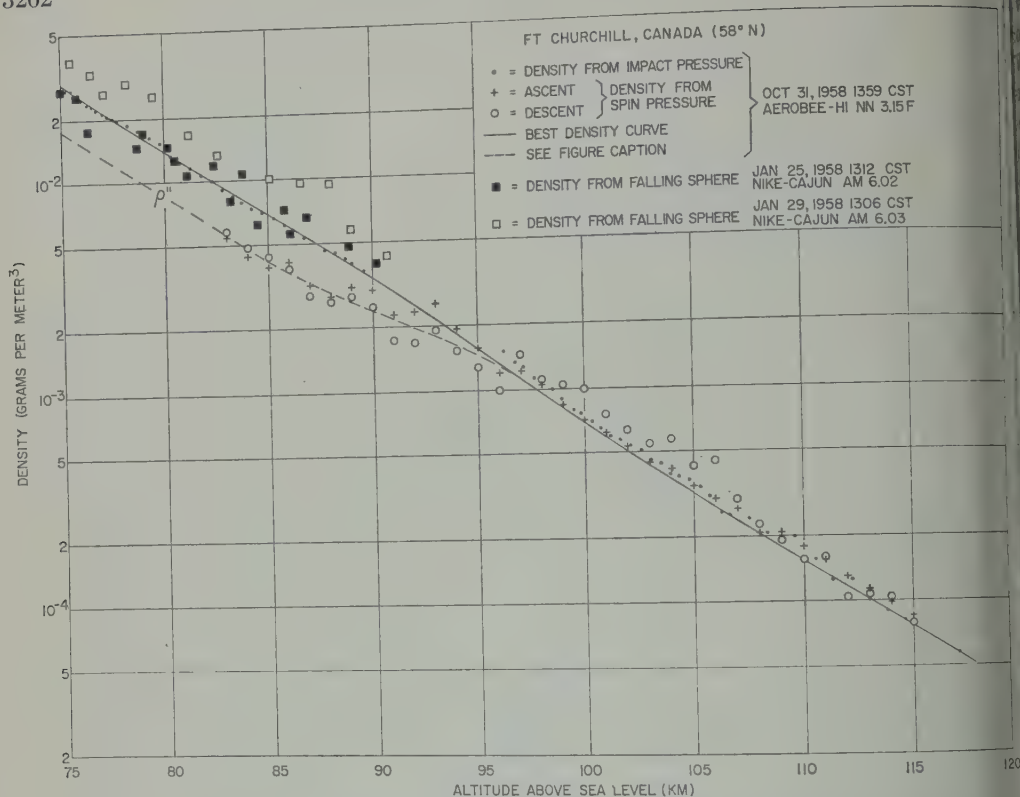


Fig. 8. Comparison of densities computed from impact pressure and from spin pressure. The dashed line was obtained when the same equation used with the spin pressure modulation for computing density in the region of free molecular flow was also used in the transition, slip, and continuum flow regions.

spin pressure maximum is determined (see Fig. 5a) by the direction of V_P , the component normal to the rocket's longitudinal axis of the rocket wind V_T . In the presence of an atmospheric wind W , the phase of the spin pressure maximum is shifted by angle δ and is determined from the direction of V_N , the component normal to the rocket's longitudinal axis of the total wind vector V . Angle δ is computed from the spin pressure measurements and from the rocket trajectory and aspect data.

Two methods were used for computing the horizontal atmospheric wind. The first method was essentially the same as that suggested by Horowitz and LaGow [1957]. In this method the measured quantities are the vector V_{Pa} and the angle δ_a at altitude h on ascent and the corresponding vector V_{Pd} and angle δ_d at the same altitude on descent. The method was graphical, however, and use was made of the simplifying

assumptions (1) that W was horizontal and equal to W_H and (2) that V_{Pa} , V_{Pd} , δ_a , and δ_d all lay in the horizontal plane. Under these circumstances there is usually a unique graphical solution for W as shown in Figures 5b and 5c. The conditions for which a unique solution is not possible are rare and are not likely to cause a large gap in the computed wind profile. Because the rocket's longitudinal axis was never more than 25° from the zenith during the wind measurements, the maximum wind-speed computation error introduced by the simplifying assumptions was less than 10 per cent. The solid line of Figure 6 shows the wind profile computed by the above method.

The dashed curve of Figure 6 presents the wind profile obtained by the second method for computing winds, a method in which only ascent data are used. The horizontal wind W_H was computed graphically by using V_{Pa} , δ_a , and

the assumptions (1) that W was horizontal and equal to W_H and (2) that V_{Pa} were horizontal. The magnitude of V_{Na} computed from (17) by using the measured ΔP_{sa} from Figure 7 and the density ρ from the impact pressure. Since the density from 98 to 114 km was parallel to the Rocket Panel density [1953], winds from 114 to 134 km were computed by using an extrapolated density was parallel to the Rocket Panel density in that region. The dotted-line extension from 114 km in Figure 6 was computed by the same method, but ρ in (17) was replaced with 'displaced' density ρ'' . This displaced density was obtained by downward extrapolation of the ascent pressure density data (see the following section) and Fig. 8) and was parallel to the density computed from the impact pressure.

The wind profile computed by the second method was judged to be more accurate than the profile computed by the first method. It should be noted that the agreement between the density computed by the two methods indicates there was no major difference between the velocity profiles in the region of ascent and descent. The region of ascent and descent was of the order of 90 km. This evidence, however, that wind profiles at the same altitude on ascent and descent may be different. *Greenhow and Neufeld* using radio observation of meteor trails in the region from 80 to 100 km, found wind profiles having vertical scales of 7 km and a horizontal extent of the order of 150 km. Like *Blamont* of the University of Paris, in his lithium-vapor-trail rocket experiments, observed photographs indicating periods of considerable atmospheric turbulence in this same

region. In free molecular flow, ambient density was also obtained from the amplitude of the spin-pressure modulation envelope using (10) with

$$P'(S) - P'(-S) = \Delta P_s$$

Equation 12 then becomes

$$\rho = \Delta P_s / (\sqrt{\pi} u' |V| \sin \gamma) \quad (16)$$

$$= \Delta P_s / (\sqrt{\pi} u' V_N) \quad (17)$$

when the spin chamber orifice is turned 90°

TABLE 3. Average Separation of the Measured Pressure and Density Curves for NN3.15F from the Computed Curves

Altitude, km	Average Separation of Measured Pressure Curve, %	Average Separation of Measured Density Curve, %
20-30	+1.5	
30-40	<0.5	<0.5
40-50	<0.5	-1.5
50-60	-2.0	-2.5
60-70	-4.0	-2.5
70-80	-3.0	-1.5
80-85	-3.0	-0.6
85-90	*	-0.5
90-98	*	*
98-100	-7.0	3.7
100-110		3.5

*Regions in which aerodynamic corrections to the static pressure and to the impact pressure were not available.

from the impact chamber orifice, γ as used in (16) is the effective angle of attack of the rocket and not the complement of the effective angle of attack as was the case in its use in (12).

Before V or V_N can be computed the atmospheric wind vector must be known. For the present computation of ρ from (17), the atmospheric wind vector was assumed to be horizontal, and an approximate value for V_N was obtained during that graphical solution in which both the ascent and descent data for determining the horizontal wind were used. Figure 8 shows the results of this density computation. The considerable scatter in the density is due to the fact that for this specific flight the computation geometry was relatively favorable for computing winds and poor for computing density. Values of V_N were often small compared with V_P and W_H , and thus small errors in these vectors and in the phase δ caused large errors in V_N . Density computed from the impact pressure was judged to be more accurate for this flight than density computed from the spin pressure. Maximum separation of the two density curves was about 10 per cent from 98 to 115 km, and no systematic separation of the two curves was observed. 'Displaced' density ρ'' is the value for the density obtained by extending the use of (17) into the region of continuum flow and is thus not a true density. It was possible, however, to

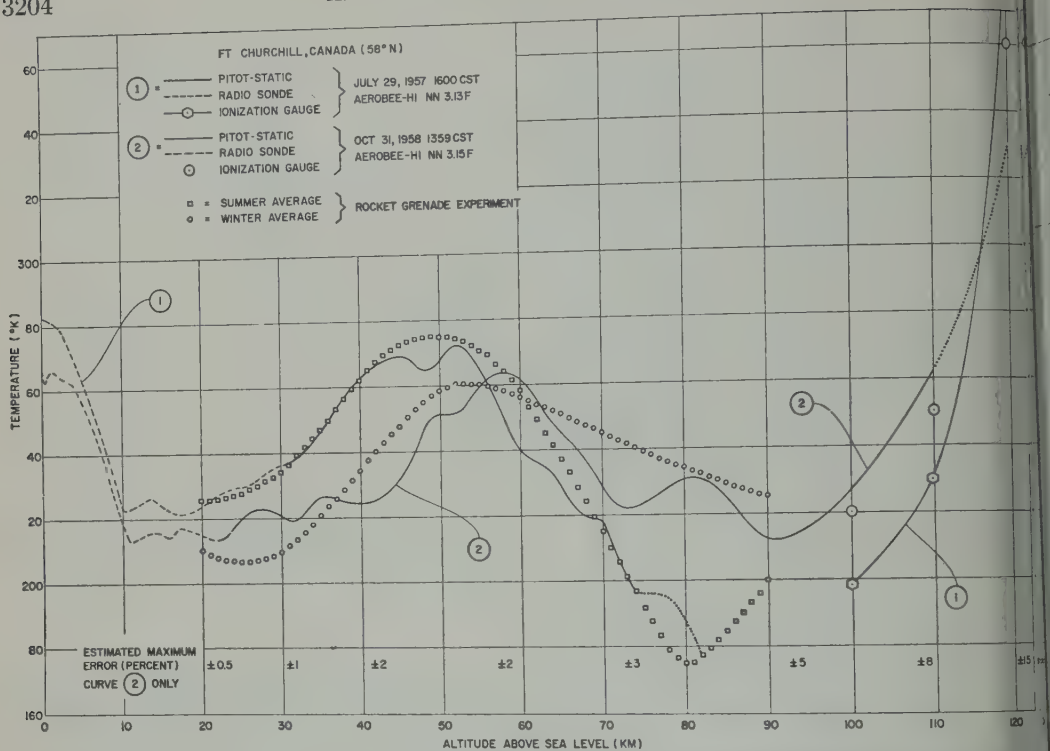


Fig. 9. Comparison of pitot-static temperature measurements with rocket grenade and ionization gage temperature measurements.

use ρ'' for determining winds in the manner described in the preceding section.

Pressure. Spin pressures were also used for computing ambient pressure. In continuum flow, ambient pressure points occur at approximately 30° on either side of the spin-pressure maximum [Hartley, 1954; Allen and Perkins, 1951; Goldstein, 1943]. The average separation of ambient pressure curves determined from static pressure and spin pressure was less than 4 per cent between 50 and 70 km, and no systematic separation of the two curves was observed. Because of the large number of spin cycles in the region between 50 to 70 km, spin pressure data are not shown in Figure 3.

TEMPERATURE COMPUTATIONS

Ambient temperature was computed from:

(1) the ratio of pressure to density as given by the equation of state

$$T = MP/R\rho \quad (18)$$

(2) the pressure profile using

$$T = -MgR^{-1} dh/d(\ln P) \quad (19)$$

where g is the acceleration due to gravity, and (3) the density profile using

$$\rho = \rho_1 \frac{T_1 M}{T M_1} \exp \left[\frac{-1}{R} \int_{h_1}^h \frac{Mg}{T} dh \right] \quad (20)$$

Equations 19 and 20 were placed in the convenient forms,

$$\bar{T} \approx MgR^{-1} \Delta h_i / \ln (P_i / P_{i-1}) \quad (21)$$

and

$$\rho_i / \rho_{i-1} \approx (T_{i-1} / T_i) \cdot \exp [-2MgR^{-1} \Delta h_i / (T_i + T_{i-1})] \quad (22)$$

The procedure used for determining a temperature profile was as follows. Radiosonde temperature T_1 , pressure P_1 , and compute density ρ_1 were used as initial reference point. At each succeeding altitude $h_i = h_{i-1} + 2$ km values of P_i and ρ_i were selected so that the values of T_i computed from (18) and (22) were

TABLE 4. Radiosonde Data for Fort Churchill, Canada

Balloon Flights								
Nov. 11, 1956, 1051 CST			July 29, 1957, 1610 CST			Oct. 31, 1958, 1150 CST		
P , mm Hg	ρ , g/m ³	T , °K	P , mm Hg	ρ , g/m ³	T , °K	P , mm Hg	ρ , g/m ³	T , °K
759	1375	256.0	755	1240	282.5	767	1320	267.0
583	1222	259.0	595	986	280.0	592	1045	265.0
447	847	245.0	465	802	269.0	458	813	261.5
336	672	232.5	357	649	255.5	349	642	251.0
248	529	217.5	270	520	241.0	265	517	236.0
182	387	218.5	202	421	223.0	197	414	219.5
132	279	219.5	149	308	224.0	144	312	213.5
96.5	204	220.0	110	226	225.5	105	224	215.5
70.5	149	219.0	81.0	169	222.5	75.7	162	214.0
52.0	109	221.0	59.8	126	221.5	54.8	118	217.0
38.2	80.4	220.5	43.7	90.9	223.5	39.2	85.0	215.0
28.2	59.2	221.0	32.3	66.0	227.0	27.9	60.7	213.5
21.0	43.7	223.0	24.0	48.6	229.0			
15.5	32.5	221.0	17.9	36.2	229.5			
11.5	24.5	217.5	13.5	26.8	234.0			
8.35	17.6	220.0	10.3	20.2	236.5			
6.15	12.7	223.5						

me and so that the temperature computed (21) lay on the T_i temperature profile between T_{i-1} and T_i . The resulting computed profiles for T_i , P_i , and ρ_i were not used, and a number of these profile sets were discarded. To select the particular T_i profile that most closely approached a true temperature profile, it was required that the computed P_i and ρ_i profiles achieve the best possible fit with the measured P and ρ profiles along the region from 22 to 100 km. An important part of this requirement of a best possible fit was that the slopes and slope changes of the computed density and pressure profiles agree with the slopes and slope changes of the measured density and pressure profiles. In the event that none of the computed sets were satisfactory, new sets were computed by using many different values of T_1 and P_1 until finally a satisfactory relation was obtained between the computed and measured profiles. Above 100 km no ambient pressure measurements were made, and the temperature was computed from the density data alone. The computed density for the region below 100 km joined smoothly with the measured density above 100 km, and this profile was used in (22) to compute the temperature.

The average separation of the measured pressure and density curves from the computed curves is presented in Table 3. Those separations having negative signs were believed to be partially due to insufficient hysteresis correction to the gages, with the exception that the error in the ambient pressure at 98 to 100 km was believed to be due both to gage errors and to error in measuring the effective angle of attack of the rocket. If Table 3 is compared with Table 2, it will be noted that correction of the measured data for the effect of the fictitious horizontal winds in the region from 20 to 75 km results in better agreement between the measured and computed density profiles.

An alternative to the use of (22) for computing temperature is the use of

$$T = \frac{\rho_1 M}{\rho M_1} \left[T_1 - \frac{M_1}{\rho_1 R} \int_{h_1}^h \rho g dh \right] \quad (23)$$

It can be shown that when the computation for the temperature is begun at a low altitude and is continued upward so that $h_i > h_1$, the value of T_i as obtained from (22) and (23) is sensitive to error in the initial temperature T_1 , to error in ρ_1 , and to error in the density profile in the region of ρ_1 ; it is relatively insensitive to error in the density profile in the region just

TABLE 5. Computed Atmospheric Structure for Fort Churchill, Canada

$h,$ km	NN3.12F			Rocket NN3.13F			NN3.15F		
	Nov. 11, 1956, 1100 CST			July 29, 1957, 1600 CST			Oct. 31, 1958, 1359 CST		
	$P,$ mm Hg	$\rho,$ g/m ³	$T,$ °K	$P,$ mm Hg	$\rho,$ g/m ³	$T,$ °K	$P,$ mm Hg	$\rho,$ g/m ³	$T,$ °K
18	51.0								
20	37.5		214.0		91.2				
22		81.1			66.6				
24		59.5			49.0		20.3	43.4	217.0
26		43.7			36.4		14.9	31.2	222.0
28		32.1			26.7		11.1	23.1	223.0
30		23.6		10.1	19.8	235.5	8.18	17.2	220.0
32		17.2		7.57	14.7	239.0	6.04	12.7	220.0
34		12.6		5.68	10.9	242.0	4.47	9.20	225.0
36		9.24		4.33	8.04	249.5	3.30	6.75	227.0
38				3.31	5.99	256.5	2.45	5.02	226.0
40				2.54	4.50	262.0	1.81	3.72	225.0
42				1.97	3.44	266.0	1.34	2.73	227.0
44				1.54	2.65	268.0	9.96(10) ⁻¹	2.00	231.0
46				1.20	2.06	268.5	7.42	1.46	236.0
48				9.30(10) ⁻¹	1.63	264.5	5.62	1.05	248.0
50				7.20	1.25	267.0	4.30	7.90(10) ⁻¹	252.0
52				5.63	9.58(10) ⁻¹	272.5	3.31	6.07	253.0
54				4.39	7.60	268.0	2.53	4.55	258.0
56				3.41	6.07	260.5	1.95	3.44	262.0
58				2.61	4.85	250.5	1.52	2.67	264.0
60				1.99	3.85	239.5	1.17	2.07	262.0
62				1.51	2.98	235.5	9.05(10) ⁻²	1.64	256.0
64				1.14	2.27	232.5	6.92	1.29	249.0
66				8.53(10) ⁻²	1.77	223.5	5.25	1.00	243.0
68				6.31	1.34	218.5	3.99	7.80(10) ⁻²	237.0
70				4.65	9.84(10) ⁻²	219.0	2.99	6.05	229.0
72				3.41	7.64	207.0	2.23	4.63	223.0
74				2.43	5.71	197.0	1.66	3.45	223.0
76					4.05		1.22	2.51	225.0
78					2.90		9.16(10) ⁻³	1.86	228.0
80					2.15		6.79	1.36	231.0
82					1.57		5.11	1.02	231.0
84							3.84	7.79(10) ⁻³	228.0
86							2.93	5.97	222.0
88							2.11	4.53	216.0
90							1.55	3.38	212.0
92							1.13	2.49	212.0
94							8.28(10) ⁻⁴	1.80	214.0
96							6.10	1.30	217.0
98							4.52	9.48(10) ⁻⁴	221.0
100							3.37	6.86	227.0
105							1.68	3.20	245.0
110							8.7(10) ⁻⁵	1.54	262.0

preceding ρ_i . The opposite condition exists when temperature computation with (22) or (23) is begun at high altitudes and is continued downward so that $h_i < h_1$. The error in T_i then becomes progressively more independent of the

error in T_1 and of the error in the density profile in the region of ρ_1 , but it is almost wholly dependent upon the error in the density profile in the region just preceding ρ_i .

For the second flight, NN3.13F, the tem

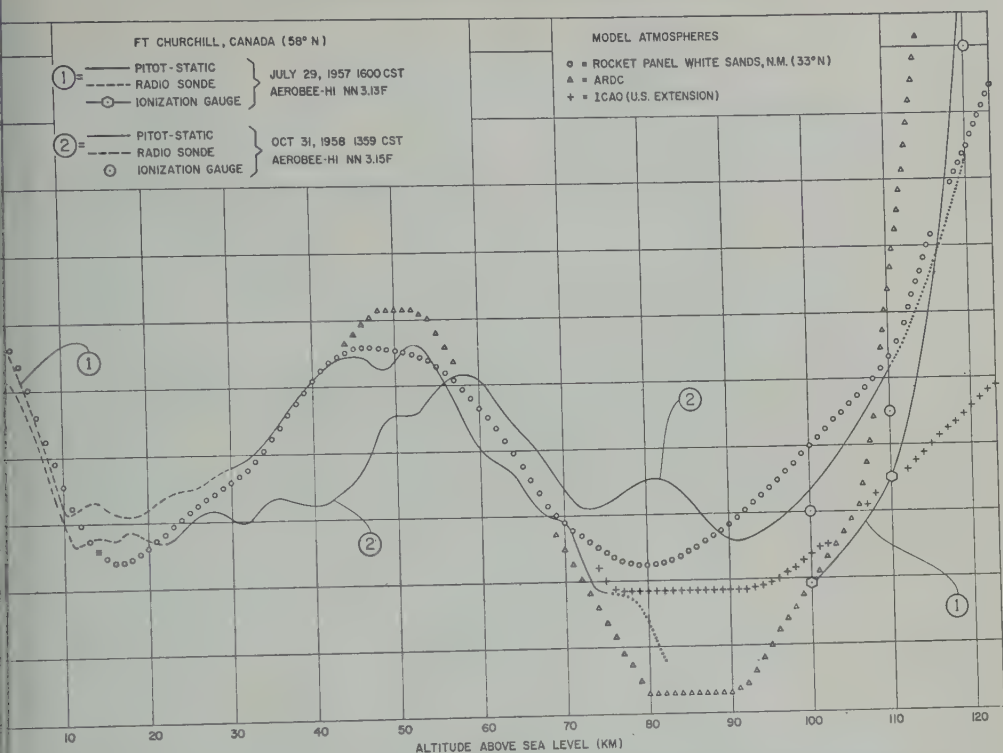


Fig. 10. Comparison of pitot-static temperature measurements with various model atmospheres.

ure at 82 km was taken from the summer-
 ge profile (see Fig. 9) obtained from the
 grenade experiment [Bandeem, Griffith,
 erg, and Stroud, 1959], and temperature
 pressure profiles were computed from the
 measured density by means of (23) and (18).
 ent pressure was measured over the range
 56 to 74 km and was used in conjunction
 the measured density to obtain the best
 able temperature, density, and pressure
 es in that region. The measured pressure
 flected the computed pressure, and the
 um separation of the two curves was
 cent.

al temperature data for all three flights
 presented in Tables 4 and 5 and Figures
 8, 9, and 10.

MEAN MOLECULAR MASS

ree values of mean molecular mass entered
 our free molecular flow equations: M_0 , the
 level value; M , the ambient value; and M' ,
 value in the gage and gage chamber. In all
 computations we have used M_0 to replace

M and M' . Thus, true values of density, pres-
 sure, and temperature must be obtained by the
 multiplications; [equation 13] $\times (M_0/M')^{1/2}$,
 [equations 15 and 17] $\times (M'/M)^{1/2}$, [equation
 18] $\times (M'M^{1/2}/M_0^{3/2})$, and [equations 21, 22,
 and 23] $\times (M/M_0)$. Meadows and Townsend
 [1960] have determined that at Fort Churchill
 at 100 km, $M = M_0$, and at 120 km, $M = 0.99$
 M_0 . On the basis of these findings we assumed
 that all the above multiplication factors were
 unity.

DISCUSSION OF RESULTS

Temperature. Temperature data are in good
 agreement with the summer-average and winter-
 average profiles determined by the grenade
 method, with the winter-average profile deter-
 mined by the falling-sphere method [Jones,
 Peterson, Schaefer, and Schulte, 1959], and with
 ionization-gage measurements [Horowitz and
 LaGow, 1958; Horowitz, LaGow, and Giuliani,
 1959].

In Figure 10 the measurements are compared
 with the model atmospheres of the Rocket

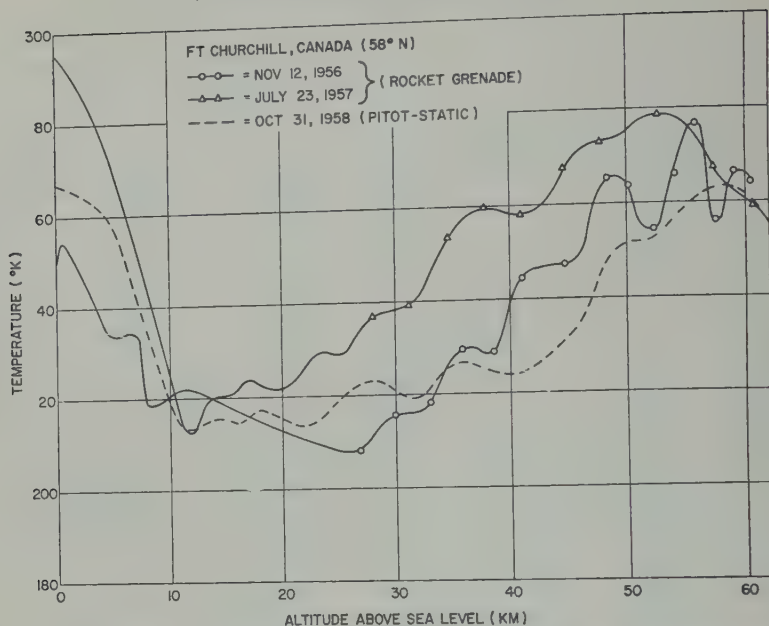


Fig. 11. A comparison of temperature fluctuations obtained by the pitot-static and grenade methods.

Panel, the ARDC [Minzner, Champion, and Pond, 1959], and the U. S. Extension to the ICAO [Minzner, Ripley, and Condron, 1958].

The small temperature minimum at 49 km for the summer-day flight and the temperature maximum at 82 km for the fall-day flight have both been noted previously [Bandein, Griffith, Nordberg, and Stroud, 1959].

The fall-day temperature profile shows minor temperature maxima at 28, 35, and 50 km. If an average temperature curve is drawn through the region from 20 to 60 km, the temperature maxima and minima represent temperature excursions of about ± 1 per cent. These temperature maxima exist despite the fact that the data used for computing the temperature profile was taken from smooth curves drawn through the impact and static pressure data points and despite the smoothing brought about by the use of 2-km altitude intervals in the temperature computation. The actual process of the successive approximation of the computed curves of pressure and density to the measured curves also led to some smoothing. True peak-to-trough amplitudes are perhaps larger than they have been depicted.

It was not possible to attribute the existence of these maxima to gage errors, to rocket motions,

or to the effects of atmospheric winds. As an example, the trough-to-peak amplitude of the temperature curve from 32 to 36 km is about 5.3°K . This change occurred in a region where the impact pressure gage had a relative error of less than 1 per cent over a 4-km altitude interval, and this error corresponds to a temperature change of less than 2.2°K . In the same region, the maximum possible effect of rocket motion on the impact pressure would cause a temperature change of less than 0.4°K . Furthermore, the 8-km altitude interval over which the temperature maximum exists would imply a 4.5-sec rocket oscillation period at a time when the actual rocket oscillation period was less than 2.5 sec. To produce the same effect on the impact pressure as a 5.3°K atmospheric temperature change would require two consecutive horizontal wind speed changes of 145 m/sec in an 8-km altitude interval or two consecutive 180° changes of a 72 m/sec wind in an 8-km altitude interval, and in both cases the azimuths of the winds would have to have a particular angular relation to the azimuthal direction of the rocket's longitudinal axis. None of the above conditions appear in published atmospheric wind profiles. The maximum speed given by the fictitious wind profiles at 35 km was 57 m/sec.

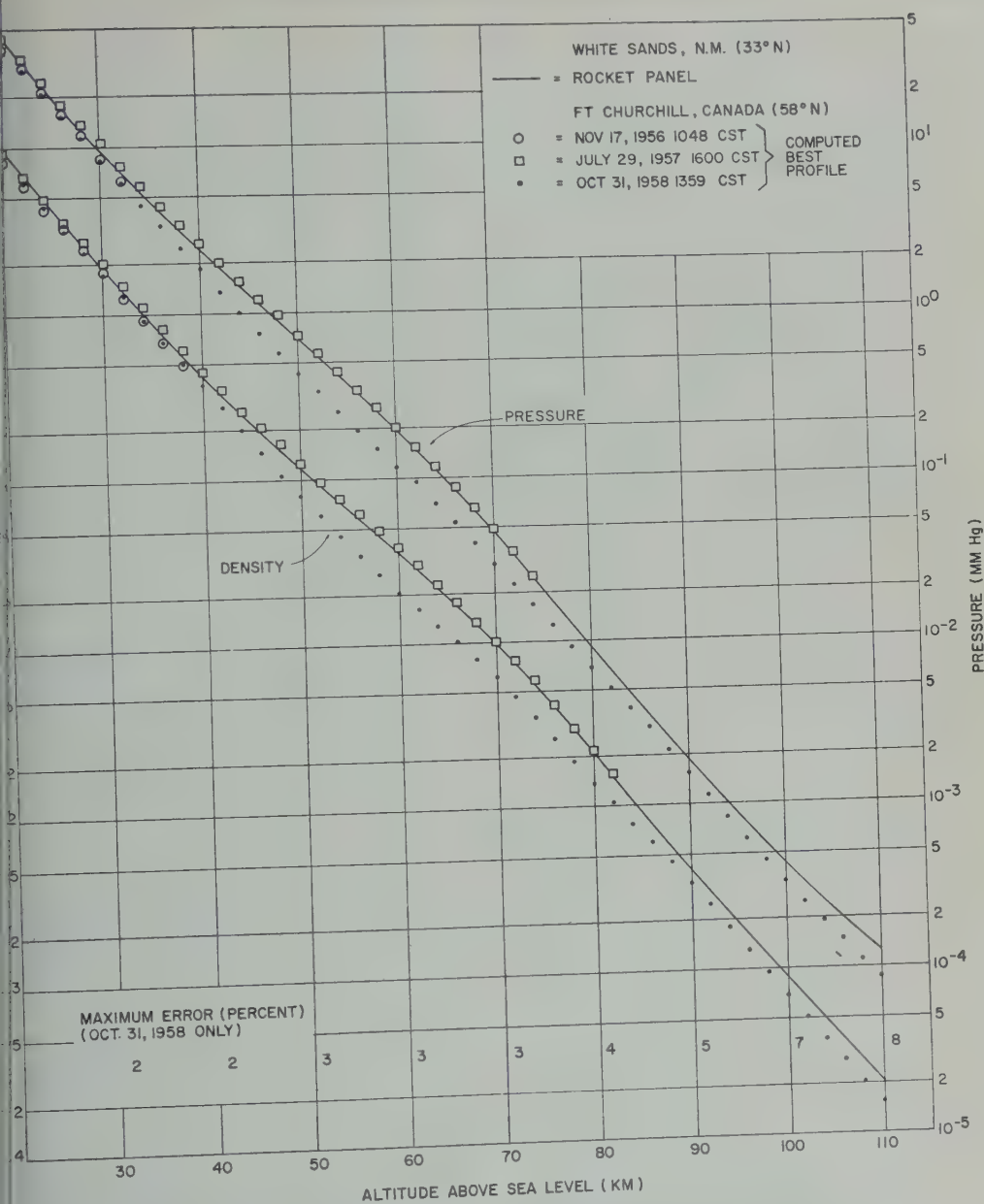


Fig. 12. Final pressure and density values for the three pitot-static tube flights.

an inspection of the temperature profiles obtained by the grenade method and by the S. S. R. [Shvidkovski, 1958] reveals that these temperature variations in the region between 2 and 50 km are more common than are smooth files. At this time, however, the variations do appear to have any obvious relation to altitude or season. Figure 11 depicts these

temperature variations on our fall-day flight and for fall and summer flights on which grenades were used. The curves drawn through the grenade data points are ours. A further discussion of these maxima and also the maxima in the measured wind profile (see Fig. 6) will be presented in a subsequent paper.

Density and pressure. In the region between

20 and 30 km the validity of the density measurements is indicated by their excellent agreement with the density computed from the radiosonde measurements. In the region between 30 and 85 km the validity of the measurements is suggested by their consistency as indicated by the close agreement between the measured and computed values of pressure and density. Additional evidence of the validity of the high-altitude density measurements is contained in the wind-tunnel measurement [Sherman, 1953] of impact pressures made with a new isentropic-flow supersonic nozzle which did not use pressure probes for calibration of nozzle flow. Sherman's work demonstrated, in effect, that true ambient density may be computed accurately from the measured impact pressure under conditions similar to those obtained in the use of the NRL pitot-static tube in the region between 69 and 84 km.

Final density and pressure profiles for the flights with the pitot-static tube are shown in Figure 12. In Figure 8 there is a limited comparison of the density for the fall-day flight with the densities obtained from two winter-day falling-sphere flights [Jones, Peterson, Schaefer, and Schulte, 1959].

Winds. The counterclockwise rotation of the wind at 85 km from west to east is in accord with the references cited by Murgatroyd [1957]. Wind direction at 82 km is toward the southwest and is in accord with the observed direction of motion of noctilucent clouds at this altitude. A profile of the magnitude of the vertical wind shear $\sigma = |\Delta W/\Delta h|$ was computed; its largest value was 90 m/sec per km at 112.5 km, and the wind direction was 70° . This largest σ occurred at the same altitude as the largest change in wind speed, $+60$ m/sec per km, and the largest azimuthal rotation rate, 21° per km clockwise.

General. The data obtained from the pitot-static tube flights indicate that the use of impact-pressure measurements to determine density is a promising method for synoptic measurements of atmospheric structure in the region from 20 to 80 km. The method has the following advantages: (1) It is possible to make continuous measurements from 20 to 80 km by means of available instruments which are simple, light in weight, and small. (2) For the most recently developed pressure sensors, the impact pressure measurement error at 80 km will be less than 2 per cent.

(3) Impact pressure measurements are not sensitive to rocket spin; thus spin may be used to reduce the maximum rocket angle of attack to the extent that no angle-of-attack correction to the impact pressure is required. (4) Vertical winds are small and have a small effect on the impact pressure. The effect on the impact pressure of horizontal winds can be kept small by means of a nearly vertical ascent trajectory. (5) High-speed machine computation of density based on equation 1 is technically feasible. The term $0.066P$ may be obtained from a standard atmosphere, since at high Mach numbers the variations in this term are of fourth order in the numerator of (1). With reasonable care in trajectory selection the rocket's departure speed can be substituted without appreciable error for $|V_T|$, and thus the term $|V_T|^2$ can be computed instantaneously from single station Doppler tracking equipment. Instantaneous rocket altitude can be obtained without appreciable error by combining the slant range, obtained from a high-speed integration of the Doppler-obtained departure speed, with simultaneous radar measurement of the angle between the zenith and the rocket's direction from the launching site.

The results from the flights also indicate that the measurement of pressure by means of a multiple-orifice static-pressure chamber is a promising method for determining the magnitude of the combined high-altitude semidiurnal and diurnal atmospheric pressure changes [Wilkes, 1949]. It has been suggested that the pressure at 40 km has its maximum value at 1000 local time and its minimum value at 1600 and that the pressure change may be as large as 10 per cent of the average pressure at that altitude. A second maximum and minimum may occur at 2200 and 0400, respectively, but the pressure change is thought to be small. In the use of a supersonic multiple-orifice static-pressure probe for the measurement of these pressure changes the following conditions hold: (1) Static-pressure measurements are not sensitive to rocket spin, and thus spin may be used to reduce the maximum rocket angles of attack to the extent that no angle-of-attack correction to the static pressure is required. (2) Vertical and horizontal atmospheric winds have a negligible effect on the measured static pressure. (3) No correction for static-pressure deficiency is required to be applied to the static-pressure measurements at

and 1600, since the difference between the static-pressure measurements is the desired. The differences in the magnitudes of the density corrections at 1000 and 1600 are likely to approach third order when compared with the magnitude of the differences in the respective pressures. (4) Stable systematic errors in calibration, gage output, telemeter presentation, trajectory measurement, and velocity measurement are likely to be considerably reduced, since only relative accuracy, not absolute accuracy, is required for the flights at 1000 and 1600.

Acknowledgments. We wish to express our appreciation to R. Horowitz for discussions concerning the problems encountered in making measurements in free molecular flow; to Dr. F. S. D. of ONR for his assistance with the dynamic problems; to M. Decker for his participation in and support of the experiment; to S. Moore, Kaslow, Gilbert, and Morris of Decker Corporation for design and fabrication of pressure gages and electronic components; to L. Secretan, A. Flanick, and W. Simms for work on getters and on the development of pressure-calibration system. Finally, we wish to express special thanks to H. Benton for his assistance over a 3-year period with the electronic and system design and with the flight preparation of the pitot-static tubes.

REFERENCES

- Research Staff, Equations, tables, and charts for compressible flow, *NACA Rept. 1135*, 1953.
- North, J. E., D. F. Fox, and H. E. LaGow, Measurement of upper-atmosphere structure by means of the pitot-static tube, *NASA Tech. Note D-670*, February 1961.
- North, J., and H. E. LaGow, Vacuum gage chamber response-time, *Rev. Sci. Instr.*, **27**(8), 1464-1465, 1956.
- North, H. J., and E. W. Perkins, A study of effects of viscosity on flow over slender inclined bodies in revolution, *NACA Rept. 1048*, 1951.
- Seiden, W. R., R. M. Griffith, W. Nordberg, and G. Stroud, Measurement of temperatures, densities, pressures, and winds over Fort Churchill, Canada, by Means of the rocket made experiment, *Army Signal Research and Development Lab. Tech. Rept., USASRD 2076*, November 1959.
- Simms, P. L. and S. A. Schaaf, The impact of density, in *Physical Measurements in Gas Dynamics and Combustion*, vol. 9 of *High Speed Aerodynamics and Jet Propulsion*, Princeton University Press, Princeton, N. J., pp. 111-123, 1954.
- Decker Aviation Corp., Bala-Cynwyd, Pa., *Final Rept.*, contract NOnr-1881-(00), amended, Sept. 1957.
- Decker Corp., Bala-Cynwyd, Pa., *Final Rept.*, contract NOnr 3000(00)(x), March 1961.
- Flanick, A. P. and J. E. Ainsworth, Thermistor pressure gauge design, *Rev. Sci. Instr.*, **32**(3), 356-358, 1961.
- Flanick, A. P., and J. E. Ainsworth, Vacuum gauge calibration system (10^{-2} to 10^1 mm Hg), *Rev. Sci. Instr.*, **32**(4), 408-410, 1961.
- Goldstein, S., *Modern Developments in Fluid Dynamics*, vol. 2, p. 422, Clarendon Press, Oxford, 1943.
- Greenhow, J. S., and E. L. Neufeld, Measurements of turbulence in the upper atmosphere, *Proc. Phys. Soc. London*, **74**, part 1, 1-10, 1959.
- Hartley, R. M., Wind-tunnel tests of two Deacon rocket models at subsonic and supersonic airspeeds, *David Taylor Model Basin Aero Rept.*, 876, November 1954.
- Havens, R. J., R. T. Koll, and H. E. LaGow, The pressure, density, and temperature of the earth's atmosphere to 160 km, *J. Geophys. Research*, **57**(1), 59-72, 1952.
- Hill, J. A. F., J. R. Baron, and L. H. Schindel, Mach number measurements in high speed wind tunnels, *Mass. Inst. Technol. Naval Supersonic Lab. Tech. Rept. 145*, January 1956.
- Horowitz, R., and H. E. LaGow, A method for determining winds in the upper-atmosphere during rocket flight, *NRL Library Document 194*, 949, Washington, D. C., 1957.
- Horowitz, R., and H. E. LaGow, Summer day auroral-zone atmospheric-structure measurement from 100-210 km, *J. Geophys. Research*, **63**(4), 757-773, 1958.
- Horowitz, R., H. E. LaGow, and J. F. Giuliani, Fall-day auroral-zone atmospheric structure measurements from 100 to 188 km, *J. Geophys. Research*, **64**(12), 2287-2295, 1959.
- Jones, L. M., J. W. Peterson, E. J. Schaefer, and H. F. Schulte, Upper-air density and temperature: Some variations and an abrupt warming in the mesosphere, *J. Geophys. Research*, **64**(12), 2331-2340, 1959.
- Kupperian, J. E., Jr., and R. W. Kreplin, Optical aspect system for rockets, *Rev. Sci. Instr.*, **28**(1), 14-19, 1957.
- LaGow, H. E., and J. Ainsworth, Arctic upper-atmosphere pressure and density measurements with rockets, *J. Geophys. Research*, **61**(1), 77-92, 1956.
- Laurmann, J. A., Low density characteristics of an Aerobee-Hi pitot-static probe, *Univ. Calif. Inst. of Engineering Research, Tech. Rept. HE-150-156*, May 1958.
- Meadows, E. B., and J. W. Townsend, Jr., I.G.Y. rocket measurements of arctic composition above 100 km, in *Space Research; Proceedings of the First International Space Science Symposium*, edited by Hilde Kallman Bijl, North Holland Publishing Co., Amsterdam, 1960.

Now at the University of California.

- Minzner, R. A., W. S. Ripley, and T. P. Condrion, U. S. extension to the ICAO standard atmosphere tables and data to 300 standard geopotential kilometers, U. S. Government Printing Office, 1958.
- Minzner, R. A., K. S. W. Champion, and H. L. Pond, The ARDC model atmosphere, in *Air Force Surveys in Geophysics*, no. 115, AFCRC-TR-59-267, Air Force Cambridge Research Center, Bedford, Mass., August 1959.
- Murgatroyd, R. J., Winds and temperatures between 20 km and 100 km—a review, *Quart. J. Roy. Meteorol. Soc.*, 83(358), 417-458, 1957.
- Newell, H. E., *High Altitude Rocket Research*, Academic Press, New York, 1953.
- Sanger, E., The gas kinetics of very high flight speeds, *NACA Tech. Mem. 1270*, May 1950.
- Schaaf, S. A., and R. R. Cyr, Time constants for vacuum gage systems, *J. Appl. Phys.*, 20(9), 860-863, 1949.
- Schultz, F. V., N. W. Spencer, and A. Reifman, Atmospheric pressure and temperature measurements between altitudes of 40 and 110 kilometers, *Upper Atmosphere Rept. 2*, Engineering Research Institute, University of Michigan, Ann Arbor, July 1948.
- Sherman, F. S., New experiments on impalpable pressure interpretation in supersonic and supersonic rarefied air streams, *NACA Tech. Note 2995*, September 1953.
- Shvidkovskii, Ye. G., *Meteorological Rocket Research of the Stratosphere*, Fifth Reunion CSAG, Moscow, 1958 (U. S. National Academy of Sciences Reprint G.2).
- Tsien, H. S., Superaerodynamics, mechanics of rarefied gases, *J. Aeronaut. Sci.*, 13(12), 653-661, 1946.
- Wiener, B., Ambient pressure determination at high altitudes by use of free-molecule theory, *NACA Tech. Note 1821*, 1949.
- Wilkes, M. W., *Oscillations of the Earth's Atmosphere*, Cambridge University Press, 1949.

(Manuscript received October 10, 1960;
revised July 17, 1961.)

Total Atmospheric Ozone and Geomagnetic Activity

S. J. AHMED AND A. HALIM

*Pakistan Meteorological Department
Karachi, Pakistan*

Abstract. Both types of emissions from the sun—the wave (W) radiation and the corpuscular (P) radiation—are known to disturb the earth's upper atmosphere as well as its magnetic field. An attempt is made to examine the relationship, if any, between total atmospheric ozone and the geomagnetic activity. No significant correlation is obtained. The reason may be that short-lived bursts of W radiation have hardly any effect on the ozone amount, or that the effect is eclipsed by the unaccountable variations in L_o , and that the P radiation perhaps never penetrates down to the ozonosphere.

INTRODUCTION

son, Harrison, and Lawrence (1927) found there is a small but definite tendency for with much ozone to be associated with magnetically disturbed conditions. Malurkar found a close relationship between the full range of atmospheric ozone and geomagnetic activity. Götz [1951] gives a correlation coefficient of 0.01 between day-to-day atmospheric ozone values and 'relative sunspot numbers', which, in their turn, are known to be among the main factors affecting the earth's magnetic field, the correlation coefficient between the means of sunspot numbers and the U index of magnetic activity being as high as 0.9 [Chapman and Bartels, 1940, p. 370]. Big [1951] went on to suggest possible variations in total atmospheric ozone following ionospheric bursts of ultraviolet radiation from the sun. Fritz [1951] attempted to measure such variations during sudden ionospheric disturbances (SID's) but found no relationship. He was, however, handicapped owing to the absence of adequate data and a lack of precise knowledge regarding any change that might have taken place during the SID's in the 'extraterrestrial constant' L_o , the log-ratio of the intensities at 3110 AU and 3300 AU outside the earth's atmosphere.

Now an attempt is made to investigate the relationship, if any, between total atmospheric ozone, the 'extraterrestrial constant,' and the earth's magnetic field, should they be dependent on the same solar causes.

TOTAL OZONE AND SUNSPOT ACTIVITY

Before examining the relationship between total atmospheric ozone and the earth's magnetism, a preliminary check was made in regard to the influence of sunspot activity on the ozone amount. Table 1 lists the correlation coefficients obtained between ozone values as measured with Dobson's spectrophotometers at Quetta, New Delhi, and at Mt. Abu over different periods and sunspot relative numbers for corresponding days as reported from Zurich.

Comparatively higher though the correlations obtained with New Delhi and Mt. Abu are, they still seem far from significant. The results obtained, in general, are in good agreement with those reported by others.

TOTAL OZONE AND GEOMAGNETIC ACTIVITY OF THE SAME DAY

Although the major variations in the total amount of atmospheric ozone are known to be due chiefly to causes other than the extraterrestrial, it was thought that statistical methods might bring to light any relationship that might exist between day-to-day ozone variations and the contemporaneous geomagnetic activity. The seasonal effects were practically eliminated from (1) ozone values by considering the departures in the daily ozone values from the monthly mean values; and (2) the geomagnetic activity by using the planetary 3-hour-range indices (K_p). These indices for the 6-hour periods including times of usual ozone observation (between 0600 and 1200 GMT for Indo-Pakistan stations and

TABLE 1. Total Ozone and Sunspot Relative Numbers

Ozone at	Data for the period	n	Standard Deviation		Correlation Coefficient
			Ozone	SRN	
Quetta	October 1, 1952-January 30, 1953	100	9.2	15.9	+ .06
Quetta	January-December 1955 (clear days only)	137	19.5	27.6	+ .09
New Delhi	January 1-April 15, 1955	100	12.8	11.9	+ .28
Mt. Abu	March 1-June 15, 1955	100	8.5	17.3	+ .50

between 0900 and 1500 GMT for European stations) were considered sufficiently representative of contemporaneous geomagnetic activity. The correlation coefficients between total ozone and K_p obtained for five stations at different latitudes, with particularly larger number of observations for Quetta to eliminate the day-to-day variations caused by circulation changes, are presented in Table 2.

The periods of observations in each case, except for Quetta (April-July 1954, 1955), were arbitrarily chosen, depending upon the immediate availability of data for the different stations. For Quetta, in addition to calculating the correlation coefficients for the same period as for Mt. Abu for the sake of comparison, a special study was made for the summer months of 1954 and 1955. The reason for a comparatively better coefficient in this case may be that day-to-day variability due to air mass changes at Quetta is least during this period. Data spread over a full year (Oct. 1953-Sept. 1954) were also examined.

According to *Fisher and Yates* [1953], correlation coefficients of this order are not significant. Yet this persistently negative sign, except

for Mt. Abu, might not be altogether accidental. And, for reasons given by *Craig* [1950] any relationship between total atmospheric ozone and solar variability should not be expected to show up more significantly. On the other hand, it may be that relationship is completely absent, a conclusion which is in agreement with the results obtained by *Dobson, Harrison, and Lawrence* [1929] when they examined the mean ozone values for northwest Europe and the mean geomagnetic character for several stations in the same region.

TOTAL OZONE AND GEOMAGNETIC ACTIVITY OF THE FOLLOWING DAY

Ultraviolet radiation from the sun reaches the earth in about 8 minutes, whereas the corpuscles leaving the sun at the same time as the ultraviolet radiation are known to reach the top of the atmosphere and affect the earth's magnetic field after about 24 hours. A possibility of some better relationship between ozone values of a day and geomagnetic activity of the following day was, therefore, considered. Hence, the ozone values of a day were paired with the K_p

TABLE 2. Total Ozone and K_p of the Same Day

Station	Lat. °N	Data for the Period	n	Standard Deviation		Correlation Coefficient
				Ozone	K_p	
Tromsø	69.7	March-June 1956	103	21.89	2.89	-.067
Oxford	51.7	March-June 1956	103	23.36	2.89	-.062
Rome	41.8	March-June 1956	103	22.67	2.89	-.001
Quetta	30.2	Aug. 18-Dec. 14, 1952	100	7.77	2.89	-.003
Quetta	30.2	April-July 1954, 1955	173	7.99	1.97	-.044
Quetta	30.2	Oct. 1, 1953-Sept. 30, 1954	247	8.42	2.02	-.001
Mt. Abu	24.6	Aug. 18-Dec. 14, 1952	100	6.59	2.89	+ .060

TABLE 3. Total Ozone and K_p of the Following Day

n	n	Standard Deviation		Correlation Coefficient
		Ozone	K_p	
	103	21.89	2.99	-.117
	103	23.36	2.99	-.168
	103	22.67	2.99	+.019
	100	7.77	2.89	-.030
	173	7.99	1.91	-.057
	247	8.42	2.02	+.042
bu	100	6.59	2.89	+.103

for the 6-hour periods on the following (27 hours' difference, i.e. between 0900 to GMT for Indo-Pakistan stations and between 1200 and 1800 GMT for European stations). The results obtained are presented in Table 3.

The correlation coefficients obtained are again too low and lead us to conclude that no significant relationship exists between atmospheric ozone and geomagnetic activity.

POSSIBLE VARIATIONS IN L_o ALONG WITH GEOMAGNETIC ELEMENTS

To account for the absence of any significant relation between total ozone and the earth's magnetism the authors felt that perhaps regular variations in L_o were taking place (owing to changes in the sun) along with geomagnetic variations, thereby introducing errors in the absolute values of ozone measured with the Dobson spectrophotometer. The negative relationship between L_o and sunspot activity has been reported by Halim [1956], and is also borne out by values given by Svensson [1958] for eclipsed sun visualizing the sunspots to 'eclipse' all fractions of the photosphere. Errors arising as a result of changes in L_o

TABLE 4. Relationship between L_o and Same Day C , A_p and K_p

Magnetic Elements Paired with L_o	Number of Observations	Correlation Coefficient
C	187	-.03
A_p	187	-.03
K_p	187	-.02

TABLE 5. Relationship between L_o and Following Day C , A_p and K_p

Magnetic Elements Paired with L_o	Number of Observations	Correlation Coefficient
C	187	+.06
A_p	187	+.07
K_p	187	+.06

possibly could prevent even a significant relationship between ozone and earth's magnetism from showing up. The investigation was, therefore, diverted to the statistical evaluation of correlation, if any, between L_o values obtained at Quetta on clear and settled days of the year 1955 and magnetic elements C , A_p , and K_p for corresponding as well as following dates. Tables 4 and 5 list the results obtained.

Though the correlation coefficients are again too low, they are all negative for same day values and positive for the values obtained with those of the following day, the sign remaining the same in each of the series considered (this is to be expected because C , A_p and K_p are, in fact, different measures of the same geomagnetic activity). This, to some extent, confirms the suspected variations in L_o which may then be responsible for eclipsing any significant relationship between total ozone and geomagnetic activity.

DISCUSSION

Solar phenomena affecting the earth's magnetic field and the upper atmosphere may be classified as follows:

1. Individual flares of *ultraviolet radiation* which produce brief geomagnetic effects simultaneously with radio fade-outs.

2. The general change of *ionizing wave radiation* during the course of the sunspot cycle which governs the intensity of the solar daily variation S_q .

3. Moderate *corpuscular radiation* which produces ordinary auroras and minor magnetic disturbances and is the main factor governing the daily magnetic character figure C .

4. Intense *corpuscular radiation* which is responsible for severe magnetic storms and auroras extending outside their normal zone and is the

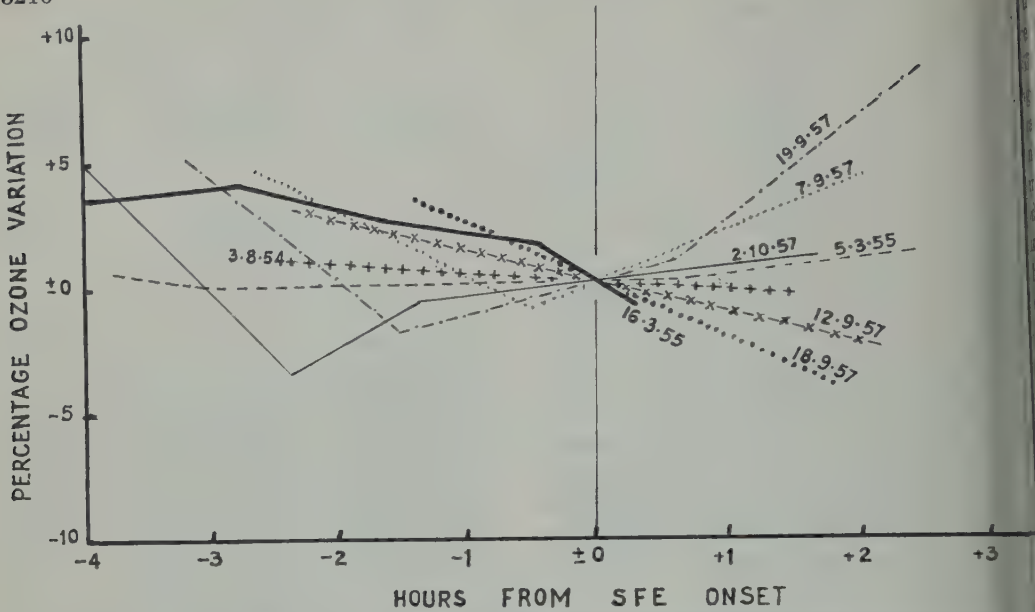


Fig. 1. Percentage variation in total atmospheric ozone from amounts at SFE onset time.

main factor affecting the U measure of magnetic activity.

Summarizing, these solar emissions can be grouped under two main types of radiation: (a) wave radiation (associated with solar flares and the bright peripheries of sunspots); and (b) corpuscular radiation (associated with solar flares and the M regions).

Wave radiation. McNish [1937] found that solar flares cause increased ionization in the D region leading to fleeting augmentations of the S_e field in the sunlit hemisphere, within 60° of the subsolar point, simultaneous with radio fade-outs. Fritz [1951] concludes that changes in total atmospheric ozone in association with SID's are small or absent provided L_o remains unaffected. A preliminary check with ozone values obtained at Quetta from observations made before and after SFE's during the last few years resulted in Figure 1, which appears to confirm the conclusions arrived at by Fritz.

Mitra [1951] speaks of the variations in the atmospheric ozone as having little, if any, association with the 11-year sunspot cycle but says that there are 27- and 15.5-day periods of variation with small amplitudes. Total ozone amount during the night is always greater (or at least never less) than that during the day.

Arguing negatively, if a fleeting phenomenon like a solar flare were capable of affecting the ozonosphere, then there ought to be a substantial change in the ozone amount during the course of the day. It does not also seem that the ultraviolet radiation associated with sunspots is influencing the increase or decrease of ozone. The authors, therefore, believe that variations in the quantum of wave radiation coming from the sun have little effect on the ozone amount of the atmosphere, at least outside the Arctic and Antarctic belts. The two polar regions have long nights (winter) and long days (summer), resulting in cumulative effects of the long absence or long influence of solar-wave radiation. If, according to the theories of Chapman, Mecke, and Wulf, light is assumed to be absorbed exponentially, or approximately so, and to give rise to ozone in proportion to (or in some function of) the energy absorbed, the ozone amounts in the polar regions may suffer conspicuous changes during the course of the year.

Corpuscular radiation. The correlation coefficient for the period 1872-1930 between sunspot numbers and the U measure of magnetic activity—a measure which is not appropriate for individual days—was + 0.869. The correlation coefficient between the sunspot number

the annual mean of the magnetic character C during the period 1906–1930 was [Chapman and Bartels, 1940, p. 394]. In contrast, Chree [1908, 1912, 1913] demonstrated the weakness of the relationship between sunspottedness of a day and the magnetic variance of the same day. From these it seems that sunspottedness, in general, has a cumulative as well as an overlapping effect on the earth's magnetic field—in the sense that it does not isolate a particular group as having a particular magnetic activity of a particular day or month. This would be clear if one is to appreciate that the rotational periods of the sun are 24.55 sidereal days at the equator to 31 days at 60° solar latitude.

Chapman and Bartels [1939] infer from the studies by Bartels and Newton [1928] of 403 magnetic storms that for most of the great storms certain limited areas of the sun's surface near the spot groups may be considered as the source of the solar streams. Not easily identifiable within the M regions of the sun by the usual physical observations are, the pronounced recurrence tendency of moderate magnetic activity. Bartels [1939] concludes that the correlation between terrestrial magnetism and traceable solar M regions.

The velocity of 700 km/sec for P radiation was assumed in the case of an emission from the sun [Fath and others, 1937]. Allen [1938] found a velocity of 750 km/sec. Milne [1926] reckoned velocities up to 1600 km/sec may be attained by matter traveling directly away from the sun. Such particles could reach the earth in hours after leaving the sun. Intervals of this order between the passage of notable sunspots across the sun's central meridian and the beginning of severe magnetic storms apparently associated with sunspots were found by Greaves and Newton [1928] and others.

The velocity of the solar corpuscles is based upon as being of the order of 1600 km/sec, Chapman and Bartels [1940, p. 809], discussing the auroral phenomena, think that the closest they can come into the earth's atmosphere there would be about 70 km above the surface.

Although there is a close relationship between auroral displays and magnetic storms on the one hand and ionospheric storms on the other, the

authors believe that even if the solar particles could have Milne's velocities of 1600 km/sec, the depths in the earth's atmosphere to which the charged particles can descend in the lower latitudes (i.e., across the magnetic lines of force) would be much less than in the auroral zones. It can, therefore, be safely inferred that the charged particles entering the fringes of the earth's atmosphere in between the auroral zones are totally diverted toward the polar regions by the earth's magnetic lines of force before they enter even the D layer, much less the ozonosphere, which extends to at most 70 km on the upper side. As for the neutral particles, their movements are apparently impeded by friction due to the earth's atmosphere to at least the same extent as, if not more than, those of the charged particles entering the atmosphere directly in the auroral zones, i.e. 70 km.

CONCLUSIONS

Visualizing the sun as the common source influencing the earth's upper atmosphere as well as its magnetic field, whether it be through the W radiation or the P radiation, this study leads us to conclude that there is little correlation between the earth's magnetic activity and changes in the total amount of atmospheric ozone.

It is not known, however, whether one would get the same result if the possible simultaneous changes in L_o are taken into account in the calculations for ozone amounts. Unfortunately it has not yet been possible to measure instantaneous values of L_o from the surface. Daily mean values of L_o seem to have little relationship with geomagnetic activity.

An incidental result of this investigation may be that the P radiation does not penetrate into the ozonosphere.

Acknowledgment. We take pleasure in expressing our thanks to Sibte Nabi Naqvi, Director of the Pakistan Meteorological Service, without whose continued guidance and encouragement this investigation would not have been possible.

REFERENCES

- Allen, C. W., *Observatory*, London, 61, 136, 1938.
- Bartels, J., *Physics of the Earth*, 8, 393, edited by J. A. Fleming, 1939.
- Chapman, S., and J. Bartels, *Geomagnetism*, I, 370, 394; II, 809, Clarendon Press, Oxford, 1940.
- Chree, C., *Phil. Trans. Roy. Soc. London A*, 208, 205, 1908.

- Chree, C., *Phil. Trans. Roy. Soc. London A*, 212, 75, 1912.
- Chree, C., *Phil. Trans. Roy. Soc. London A*, 213, 245, 1913.
- Craig, R. A., *Am. Meteorol. Soc. Monograph*, 1, 2, 37, 1950.
- Craig, R. A., *Compendium of Meteorology*, American Meteorological Society, Boston, p. 292, 1951.
- Dobson, G. M. B., D. N. Harrison, and J. Lawrence, *Proc. Roy. Soc. London A*, 114, 521, 1927.
- Dobson, G. M. B., D. N. Harrison, and J. Lawrence, *Proc. Roy. Soc. London A*, 122, 456, 1929.
- Fisher, R. A., and F. Yates, *Statistical Tables*, 4th ed., Oliver & Boyd, Edinburgh, Table VI, p. 54, 1953.
- Fritz, S., *Arch. Meteorol. Geophys. u. Bioklimatol.*, A, 4, 343, 1951.
- Götz, F. W. P., *Compendium of Meteorology*, American Meteorological Society, Boston, p. 275, 1951.
- Greaves, W. M. H., and H. W. Newton, *Monthly Notices Roy. Astron. Soc.*, 88, 5561, 1928.
- Greaves, W. M. H., and H. W. Newton, *Monthly Notices Roy. Astron. Soc.*, 89, 84, 1928.
- Halim, A., Dependence of L_o on sunspots,—Paper presented at the International Symposium on Atmospheric Ozone, Weissenau, Germany, June 1956.
- Malurkar, S. L., *Ann. di Geophys.*, 7, 2, 209, 1952.
- McMath, R. R., and others, *Publ. Astron. Soc. Pacific*, 49, 3, 1937.
- McNish, A. G., *J. Terrest. Magn.*, 42, 109, 1937.
- Mitra, S. K., *Compendium of Meteorology*, American Meteorological Society, Boston, p. 254, 1951.
- Milne, E. A., *Monthly Notices Roy. Astron. Soc. London*, 86, 459–473; 578–599, 1926.
- Svensson, B., *Arkiv Geofysik*, 2, 28, 573, 1958.

(Manuscript received April 14, 1961;
revised July 26, 1961.)

Airborne Filters for the Measurement of Atmospheric Space Charge

C. B. MOORE, B. VONNEGUT, AND F. J. MALLAHAN

*Arthur D. Little, Inc.
Cambridge, Massachusetts*

Abstract. Airplane instrumentation is described for measuring the concentration of atmospheric electric space charge by the utilization of a glass microfiber filter that collects effectively all ions and particulate matter from the air passing through it. This equipment gives space-charge-concentration values that are in good agreement with those estimated independently from the rate of change of the vertical potential gradient with altitude in clear, filtered air. Spurious indications of high concentrations of space charge were found when the apparatus was flown into clouds. This difficulty was traced to the undesirable charging of some cloud and precipitation particles that inevitably touched the entrance to the apparatus. The authors' experience leads them to doubt the validity of all airplane measurements whether made with their devices or with any devices yet reported—of the charge on cloud tops. Outside the clouds the apparatus has been found useful on many flights in the direct measurement of natural and artificial space charge.

INTRODUCTION

Electrical activity of both fair and dis-weather arises from a complicated distribution of space charge in the atmosphere. It is, therefore, of interest and importance in the study of atmospheric electricity to know how the concentration of this charge varies with time and through the depth of the atmosphere. Since it is impossible to determine the distribution from ground measurements it is desirable, if we are to understand the electric electrical process, to be able to measure space-charge concentration from an aircraft. Inasmuch as present aircraft instrumentation for measuring space-charge concentration leaves much to be desired, we have devoted some effort to the development of such equipment.

Among the various possible schemes for space-charge measurement from an aircraft, one of the most suitable appears to be the technique of Obolensky [1925] in which charge densities are determined by the measurement of the electric current that flows when the air is passed through a filter. Various investigators [for example, Brown, 1930; Gonsior, 1956] have used this type of apparatus for measuring space charge from the ground. Because it can be made simple and light in weight, such equipment is adaptable to aircraft use.

FILTRATION MEDIA

In this kind of apparatus, the filter is of critical importance. To provide a correct measurement, it must remove all ions and charged aerosol particles from the air that passes through it. A further requirement is that the filter have a sufficiently low pressure drop to allow air to pass through it at a sufficiently high volume rate of flow to give a readily measurable rate of charge capture.

Our laboratory tests showed that, although a steel-wool filter medium of the kind used by Obolensky had the advantage of a fairly low pressure drop, it was a poor filter for removing smokes, Aitken nuclei, and ions. Accordingly, we concluded that steel wool was not suitable for this type of apparatus and sought something more satisfactory.

We have found, following laboratory work performed by W. D. Crozier and M. Bloom (private communication, 1958), that a recently developed filter medium composed of glass microfibers is well suited to this type of apparatus, for it combines a small pressure drop with excellent filtration.

This medium was developed for the complete collection of radioactive particulate matter in atomic-energy installations [Smith, 1949; Smith and Surprenant, 1953; Stafford and Smith, 1951, 1952]. Known commercially as the Absolute

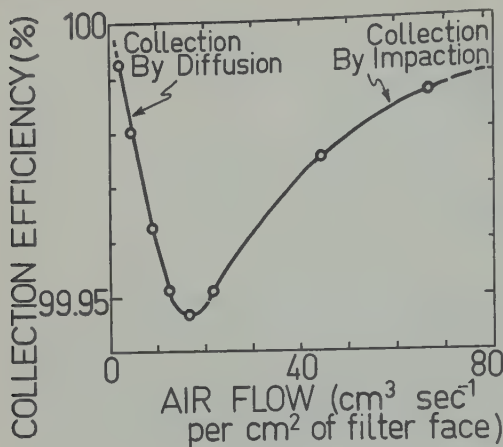


Fig. 1. Collection efficiency of Arthur D. Little glass microfiber absolute filter in removing $0.3\text{-}\mu$ particles of dioctyl phthalate dispersed in air, as function of air flow [after *Stafford and Smith*, 1952].

Filter, this filter paper is a feltlike material composed of $0.5\text{-}\mu$ glass fibers dispersed in and supported by a mat of glass fibers $3\text{ }\mu$ in diameter with a small amount of an acrylic binder and a dispersing agent. At standard conditions and with an air flow of $10\text{ cm}^3\text{ sec}^{-1}\text{ cm}^{-2}$ of face area, the paper has a pressure drop of no more than 8 mb and a collection efficiency of 99.9 per cent or more for an artificial test fog composed of dioctyl phthalate droplets $0.3\text{ }\mu$ in diameter. The collection efficiency of a sample of this medium as a function of air velocity is shown in Figure 1.

We have investigated the properties of this filter medium for the collection of atmospheric condensation nuclei, using a photometric method of evaluation [Vonnegut, 1949]. We drew laboratory air containing approximately 10^6 nuclei cm^{-3} through the filter at various speeds and measured the particle concentration in the filtered air. Usually the filter removed 99 per cent or more of the condensation nuclei from the incoming air at all flow rates. The polar conductivity of air after passage through the filter under all conditions was found to be less than 1 per cent of the conductivity of laboratory air ($\sim 10^{-14}\text{ mho m}^{-1}$).

As a comparison we measured the filtration effectiveness of a steel-wool filter that we constructed. This filter, 12 cm thick and 6 cm in

diameter, was composed of densely packed rough-edged, stainless-steel fibers 50 to 100 μ in diameter. At higher flow rates (about $10\text{ liters sec}^{-1}$), steel wool reduced the laboratory air nuclei concentration by about 20 per cent; at low flow rates (about $0.03\text{ liter sec}^{-1}$) the steel wool removed about 95 per cent of atmospheric nuclei. At the higher rates, the filter decreased the conductivity of laboratory air by about 90 per cent. This indicates that some fast ions can escape collection passing through the densely packed steel wool.

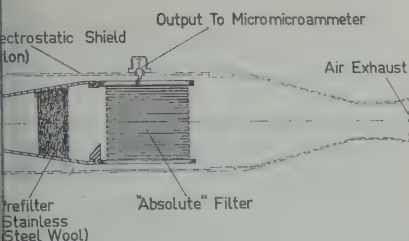
When cigarette smoke was blown into the steel wool, much of it passed through. This suggests another difficulty with a steel-wool filter. If aerosol particles underwent elastic collisions with the metal fibers and passed through without being captured, some of them might become charged by contact electrification. Their escape from the steel wool with one polarity of charge would suggest the arrival in the steel wool of charges of the opposite polarity. In view of these measurements, we believe that results obtained with steel-wool filters merit skepticism.

From the results of our laboratory studies it appears that the glass-microfiber filter medium removes essentially all the charge carriers and particulate matter in the atmosphere. It is reasonable, therefore, to conclude that the amount of charge collected by the high-efficiency filter is a fair indication of the imbalance between positive and negative charge in the air being filtered.

APPARATUS

The simplest of the two different versions of the charge-measuring apparatus we have used is shown in Figure 2. It consists of a filter cartridge mounted in a metal tube that is supported on insulators in a metal housing that serves as an electrostatic shield. This assembly is mounted on the airplane so that air at full ram pressure flows into the apparatus and through the filter.

Air intake and filter. The air intake was constructed of Teflon so that it would not be subject to point discharge under conditions of high potential gradient. A prefilter of stainless steel wool was used to protect the microfiber cartridge from collisions with insects or rain during flight. The filter cartridge, a Cambridge Corporation model 1F-20-2S, was $10 \times 10 \times$



Filter apparatus for clear-air measurement of atmospheric space charge from light

ze. The filter and the prefilter together pass 1 liter of air sec^{-1} for each millibar of drop across the combination. An ejector is mounted on the exhaust end of the electrostatic shield to produce the maximum drop across the filter. The filter assembly is mounted under the left wing of a light aircraft beneath a conventional meteorograph

cal insulation. In this apparatus it is of considerable importance to maintain the best electrical insulation, for leakage and currents can interfere seriously with results. Because Teflon (a fluorinated plastic) is less affected by humidity than most insulators, it is well suited to this work, and has been used throughout.

We have found that it is quite important in the design and assembly of the apparatus to be sure that the insulators are mechanically rigid to minimize electrical noise from piezoelectric effects and changes in capacitance. For the same reason it is desirable that the electrical cables from the apparatus and the micromicroammeter be kept as short as possible.

Flow. The air flow into this filter varied as the square of the air speed; it was not linear. If Q is the rate of air flow, then

$$1 \times 10^3 \Delta P \text{ cm}^3 \text{ sec}^{-1} \text{ mb}^{-1} \quad (\text{measured})$$

that is, the pressure drop ΔP is the difference between the ram pressure and the ejector pressure at the indicated air speed V ; ρ_a is the air density at the conditions for which the air-speed indicator is calibrated. The total pressure measured with our apparatus is estimated to be 1.8 times the ram air differential pressure

$$\Delta P = 1.8 \left(\frac{1}{2} \rho_a V^2 \right)$$

and

$$Q = 2.3 \times 10^{-3} V^2 \text{ cm sec}$$

When this filter was used, the indicated air speed had to be recorded frequently, for it varied from about 27 to 55 m sec^{-1} . In much of our work, the indicated air speed was about 30 m sec^{-1} ; for this we estimate that the filter was passing 21 liters of air sec^{-1} . At our maximum speed, about 70 liters sec^{-1} were filtered.

Charge determination. The charge collected by the filter flowed to the electrostatic shield through a low-noise RG-8A coaxial cable and a battery-operated, sensitive micromicroammeter (Keithley model 600). In turn, this last operated a d-c amplifier (Texas Instrument type DC-301) to drive a recording milliammeter (Esterline Angus). The instruments and recorder were located in the aircraft cabin. The micromicroammeter measured the potential drop across a 10^{10} -ohm resistor; the sensitivity often used was the 30-mv scale. The response time of this filter apparatus was a few seconds, about the same as the time resolution of the recorder. Use of this battery-operated measuring equipment minimized the usual difficulties due to interaction with other aircraft equipment and with the power supply.

The mean space-charge concentration N in the atmosphere through which the filter was flown was estimated by dividing the measured electrical current I from the filter by the air flow Q through the filter computed from the observed indicated air speed V , and the measured filter characteristics:

$$\bar{N} = \frac{I}{Q} = 2.8 \times 10^{21} \frac{I}{V^2} \left(\frac{\text{elementary charges}}{\text{cm sec}^2 \text{ ampere}} \right)$$

MEASUREMENTS IN CLEAR AIR

This apparatus was used successfully on 35 flights during the summer of 1960, largely in the region in central Illinois near Champaign. During these flights natural space charge was measured under a variety of different conditions and also space charge produced at the ground from an artificial source. Simultaneous atmos-

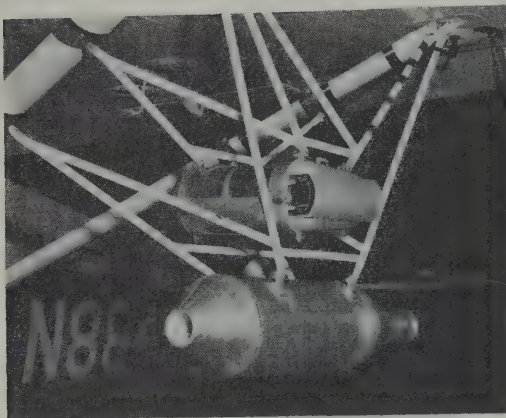


Fig. 3. Filter apparatus mounted on light plane below meteorograph.

pheric electrical measurements were also made with an airplane instrument for measuring the vertical component of the potential gradient [Vonnegut, Moore, and Mallahan, 1961]. We observed that the indications of the space-charge apparatus seemed to follow closely the vertical rate of change of the potential gradient in stratified air. During fair weather, as the aircraft rose through a haze layer in the atmosphere, an indication of negative charge was usually found at the base of the layer, where the potential gradient increased rapidly. When the aircraft surmounted the haze, positive charge was invariably indicated, accompanied by a rapid decrease in the potential gradient. An example of such a sounding is shown in Figure 4. As might be expected, the space-charge recordings usually showed a more detailed structure than did those for the potential gradient.

Since under stratified conditions the space charge measured directly always followed closely that inferred from the rate of vertical change of the potential gradient by use of Poisson's relation, we are encouraged to believe that in clear air this apparatus gives a relatively reliable indication of the true space-charge concentration in the atmosphere.

Flights through plumes of space charge produced by an artificial source (which will be described in another paper) invariably produced correct indications of the polarity of the charge emitted. Figure 5 shows examples of this and of the character of the record obtained during our flights in clear air.

The maximum current so far observed on this filter is -10^{-11} ampere. This was on Aug. 18, 1960, when we were flying in clear air at an altitude of 600 m, under the edge of a thunderstorm near Clinton, Illinois. Since our air speed was about 53 m sec^{-1} , this current indicates that we were flying through clear air in which the space charge was negative and there was a concentration of about 1000 elementary charges cm^{-3} . The horizontal extent of this volume of charge was about 3 km. During this period the vertical component of the potential gradient increased to $+50 \text{ v cm}^{-1}$. It then decreased to v cm^{-1} as we passed the charge maximum. The storm system was much too complex for us to obtain an understanding of the charge distribution around its base, but the performance of the instrument gives us hope of characterizing the charge around simpler cloud systems.

Experimental difficulties. Under certain conditions the cable connecting the collector with the micromicroammeter produced spurious indications on the recorder chart. Some of these arose from the flexing of the cable as the aircraft was maneuvered. Engine vibration also caused cable flexure, which provided a constant background 'noise' signal, usually equivalent to about 2 or 3 elementary charges cm^{-3} in the air passing through the filter.

Whenever the aircraft rose into cooler air, the thermoelectric properties of the low-noise cable produced an indication of negative charge arriving at the filter (a negative potential of 0 or 3 mv then appeared across the cable output).

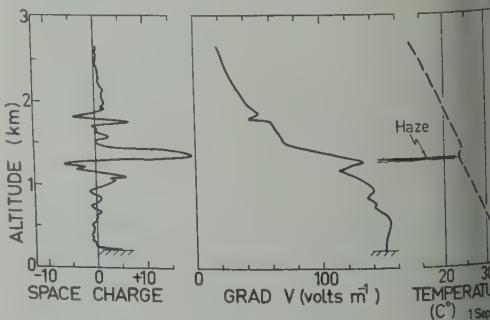
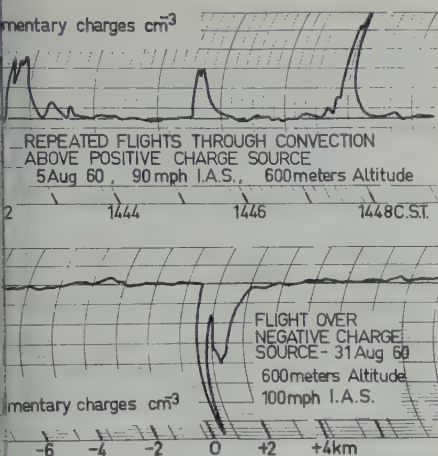


Fig. 4. Atmospheric sounding of space charge, potential gradient, and temperature on Sept. 1960, over central Illinois. It can be seen that the space-charge indications are generally proportional to the rate of change of the potential gradient with altitude.



Typical recordings of space charge in air convection over artificial charge source.

cent, a spurious indication of positive was recorded (a positive potential of 4 v appeared across the cable output on descent). The thermal effect was identified when we placed various components in a dry cold box and observed the output. We are attempting to eliminate this effect in future apparatus by shorting and thermally insulating the connecting wires. We also plan by introducing an air valve and filter to enable the operator to make independent determination of a dynamic zero. We should note that although zero drift and background noise in the apparatus were a source of annoyance, they did not interfere seriously with our measurements. After a little experience we were able to recognize these spurious indications clearly. Though they prevented us from making good measurements of very low space charge concentrations, they were negligible in comparison with the larger currents arising from space charge at inversions, haze layers, near the ground at night, as well as the artificial space charge we produced ourselves. Sometimes, the micromicroammeter produced indications of charge collection when the aircraft was in high electric fields. Since no indications were produced even though the antenna was disconnected, we believe that the RF noise caused by point discharge close to the antenna was being detected by the electrometer. The indications followed the radio static characteristic sawtooth pattern of dis-

charge and relaxation, so they were easily identified. We hope to eliminate this effect also in future instrumentation.

CHARGE MEASUREMENTS IN CLOUDS

Unquestionably the most interesting airplane space charge measurements to be made are those in clouds. We have found, as might be expected, that measurements here are far more difficult than in the clear air. The serious problem arises of bringing the cloud sample into the apparatus without changing the drop size distribution and liquid water content of the cloud or changing the charge carried by its drops.

Although our effort to make meaningful airplane measurements of space charge in clouds has thus far been generally unsuccessful, it seems worth while to describe the experiences which we have found instructive, not only in understanding the problems involved in apparatus design but also in evaluating similar measurements in clouds that other workers have reported.

Most of our attempts to measure the space charge in clouds have been made with a filter apparatus somewhat similar to the one we have just described but specifically designed for use in clouds. The design of this apparatus is shown in the drawing in Figure 6, and the way it was mounted in the nose of a P38 airplane is shown in Figure 7. It can be seen that besides being somewhat larger than the other apparatus (a Cambridge Filter Corporation type 1E-80

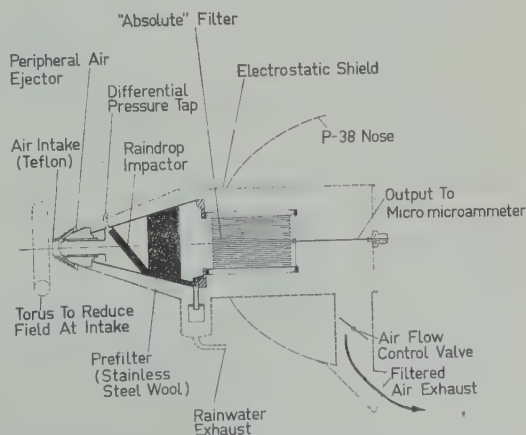


Fig. 6. Filter apparatus for measurements of atmospheric space charge in clouds and clear air.



Fig. 7. Filter apparatus installed in nose of P-38 aircraft.

20.5 × 20.5 × 28 cm filter is used in it) it has several other features intended to make it suitable for operation in clouds. The metal torus mounted ahead of the intake is installed for the purpose of reducing, in strong fields, the likelihood of corona at the intake which might cause spurious charging of cloud drops. The air flow into the apparatus was made isokinetic so that a proper sample of cloud particles might be taken.

As in the other apparatus, to protect the microfiber filter, two prefilters of stainless steel wool were placed in the conical duct to intercept incoming rain and cloud drops. Water collected in the prefilters was allowed to flow out of the bottom of the apparatus through a special 'field-free' exit so that the water would take no electrical charge with it.

We found that, whereas this apparatus, like the other, gave quite reasonable and consistent values for the space charge in clear air, it gave extremely high and quite unreasonable indications inside of even the smallest cumulus humilis clouds. We are sure that these indications were caused, not by the charge naturally carried by the cloud droplets but by the spurious charges some of them acquired when they touched the inlet as they entered the apparatus. We experimented with inlets made of different materials and found that with Teflon we obtained indications of spurious high positive charges in clouds and, with a metal intake, spurious negative charges.

We were interested in the extent to which a metal knife edge intake arrangement produced corona in strong fields. When we flew through

clear air toward an electrified cloud, concentrated space charge of one polarity was indicated; but as we passed over the top of the cloud the indicated polarity reversed. Probably the radial electric field produced by the cloud caused the aircraft, as it approached, to become so longitudinally polarized that corona discharge occurred at the knife edge, producing space charge at the intake. As the aircraft passed over the cloud top, the longitudinal polarization vanished and the indicated charge went to zero momentarily. Then, as the aircraft left the cloud, reverse polarization and corona began, producing spurious space charge of polarity opposite to that found while approaching the cloud.

COMPARISON OF OBSERVATIONS OF SPACE CHARGE IN CLOUDS

It is very interesting that our anomalous measurements within clouds are quite similar to measurements reported by Gunn [1952]. He attempted to measure the charge on cloud droplets in fair-weather cumulus by using a 'cyclone separator' as a centrifuge to collect the large droplets and an electrostatic precipitator to collect ions. He reported that the centrifuge collected droplets as small as 10 μ in diameter and he estimated that the precipitator collected all particles with an equivalent diameter of less than 0.01 μ (i.e., ion with mobilities greater than 3×10^{-8} cm sec⁻¹ per v cm⁻¹). The combination of these two instruments did not collect particles in the cloud with diameters greater than 0.01 to 10 μ , nor (probably) singly charged Aitken nuclei. Gunn noticed that the charge flowing from the centrifuge usually differed not only in polarity but slightly in magnitude from that flowing from the precipitator. He suggested that the difference represented charge on particles between 0.01 and 10 μ in diameter (the one he could not measure) since the cloud as a whole appeared externally to be electrically neutral.

As a result of his measurements, Gunn stated that in slowly developing inactive fair-weather cumuli, there might be about 250 drops cm⁻³ averaging perhaps 20 μ in diameter and each carrying an average of 32 positive electron charges. Outside each drop, he concluded, was a sheath of electrical charge that almost exactly neutralized the charge carried by the cloud

Since a field of about 4.5 v cm^{-1} would be
 ed at the droplet surface by the presence
 positive charges on the droplet, it is diffi-
 understand how such a mixture of
 drops and oppositely charged ions could
 for any time.

Though Gunn discounts this possibility, our
 nce leads us to feel that his results may
 been strongly affected by drop breakup or
 al electrification in the centrifuge. Gunn
 o experiments with his apparatus in
 neutral droplets were used to show that
 did not have occurred.

can test Gunn's interpretation of his re-
 nce our apparatus collects, in one filter
 y, essentially all fast-ion, slow-ion, cloud,
 ndrop charge carriers that pass into the
 e. Since this apparatus indicates a high
 ration of space charge in small, neutral
 and since no appreciable fraction of the
 carriers escapes the filter, we can con-
 nly that the high indications are spur-
 ing from the sampling of the cloud by
 paratus. This conclusion is reinforced by
 recovery that we could, at will, vary the
 ed polarity of the space charge by chang-
 intake walls from a dielectric to a con-
 or vice versa.

conclude that, thus far, the measurement
 rge on cloud and rain droplets by air-
 n flight is a difficult and still unsolved
 n because of the lack of a suitable means
 ecting undisturbed drops for the measure-
 of the charge they carried before sampling.
 Measurements in fair-weather clouds with
 er, as well as previous airplane measure-
 such as those of Gunn [1952] and Whit-
 [1958], give space-charge values that can-
 be believed. The considerably lower values
 ve have obtained using a balloon-borne
 y cage [Moore, Vonnegut, and Emslie,
 are undoubtedly much closer to the truth.

plan to continue our efforts to develop
 harge measuring equipment that will
 e satisfactorily in clouds. Preliminary
 ave indicated that the smaller apparatus
 on the light airplane, although not de-
 for this purpose, gives considerably more
 able measurements in clouds than the
 apparatus. This difference apparently
 from the much lower air speed through

the intake of the smaller device. Accordingly,
 it may be possible in future equipment to re-
 duce or eliminate spurious charging by further
 reduction in the air velocity.

CONCLUSIONS

Although much further work will be required
 to develop suitable airplane apparatus for meas-
 uring space charge in clouds, it appears that in
 its present form the filter apparatus that we de-
 scribe is a reliable instrument for measuring
 space charge in the clear air. The results ob-
 tained with it will be described in a forthcoming
 paper.

Acknowledgments. We wish to thank the Geo-
 physical Research Directorate of the Air Force
 Cambridge Research Center and Major T. O.
 Haig, U.S.A.F., for the support of contract AF19
 (604)1920, under which our investigation of space-
 charge filters began. We also wish to thank the
 Office of Naval Research, the Aerology Division
 of the Bureau of Weapons, and the National Sci-
 ence Foundation for supporting the development
 and flight tests of the space-charge filters reported
 here.

The success we found with the apparatus arises
 primarily from the remarkable properties of the
 glass fiber filtration medium. We wish to thank
 Walter J. Smith and his associates at Arthur D.
 Little, Inc., who developed the absolute filter, and
 Dr. William D. Crozier of the New Mexico Insti-
 tute of Mining and Technology who pioneered in
 the use of this material for space-charge measure-
 ments. We are grateful to Paul Engle of the Cam-
 bridge Filter Corporation, who supplied us with
 special cartridges to meet our needs.

We are grateful to James Cook for installing
 the large filter apparatus on his P-38 and for his
 many helpful suggestions about proper air sam-
 pling. C. F. vanThullenar and the United States
 Weather Bureau are to be thanked for the loan
 of the instrumentation on the P-38 and for their
 cooperative help. We also appreciate the assistance
 given freely by Glenn Stout, Donald Staggs, and
 J. W. Bullock of the Illinois State Water Survey
 in making the light plane available and in flight-
 testing the apparatus with us.

REFERENCES

- Brown, J. G., The relation of space charge and
 potential gradient to the diurnal system of con-
 vention in the lower atmosphere, *Terrest. Mag-
 netism and Atmospheric Elec.*, 35, 1-15, 1930.
- Gonsior, B., Messungen der Luftelektrischen
 Raum Ladungen mit Schwebstoff-Filtern, Re-
 port on the meeting on *Atmospheric Electricity
 During I.G.Y.*, Aachen, May 1956, edited by H.
 Israel, pp. 17-23, 1956.
- Gunn, R., The electrification of cloud droplets in

- nonprecipitating cumuli, *J. Meteorol.*, 9, 397-402, 1952.
- Moore, C. B., B. Vonnegut, and A. G. Emslie, *Observations of Thunderstorms in New Mexico*, Report to the Office of Naval Research under contract Nonr 1684(00), Arthur D. Little, Inc., 1959.
- Obolensky, W. N., Über Elektrische Ladungen in der Atmosphäre, *Ann. Phys. Lps.* 77, pp. 644-666, 1925.
- Smith, W. J., *Investigation of Stack Gas Filtering Requirements and Development of Suitable Filters*, Report No. 7 to Atomic Energy Commission, ALI 18, Arthur D. Little, Inc., September 30, 1949.
- Smith, W. J., and N. F. Surprenant, Properties of various filtering media for atmospheric dust sampling, *Proc. Am. Soc. Testing Materials*, 53, 1122-1133, 1953.
- Stafford, E., and W. J. Smith, Dry fibrous filters for dust free air, *Ind. Eng. Chem.*, 43, 1346, 1951.
- Stafford, E., and W. J. Smith, Dry fibrous filter for high efficiency air cleaning, *Air Pollution*, edited by Louis C. McCabe, McGraw-Hill Book Co., New York, p. 264, 1952.
- Vonnegut, B., A continuous recording condensation nuclei meter, *Proc. Natl. Air Pollution Symposium, 1st Symposium, Pasadena, California, 1949*, pp. 36-44, 1949.
- Vonnegut, B., C. B. Moore, and F. J. Mallahan, Adjustable potential-gradient-measuring apparatus for airplane use, *J. Geophys. Research*, 66, 2393-2397, 1961.
- Whitlock, W. S., *An Investigation Concerning the Electrical Properties of Fair Weather Cumulus Clouds*, Admiralty Research Laboratory, Teddington, Middlesex, 19 pp., 1958.

(Manuscript received May 1, 1961.)

Natural Radioactivity in the Atmosphere

RAMA¹ AND M. HONDA

*School of Science and Engineering
University of California, San Diego
La Jolla, California*

Abstract. The activities of cosmic-ray-produced radioisotopes P^{32} , P^{33} , and Be^7 , and of Pb^{210} , which arises from the decay of radon, have been measured in samples of dust filtered in stratospheric and tropospheric air. The short-lived activities appear to be in secular equilibrium with the respective local productions in various regions of the upper stratosphere. The activities of Pb^{210} , on the other hand, are found to be uniform throughout the stratosphere. The implications of these and the tropospheric observations are discussed.

Introduction. A number of cosmic-ray-produced short-lived isotopes have recently been found in the earth's atmosphere. The notable ones among these are P^{32} (14.5 d), P^{33} (Be⁷ (53 d), S^{35} (87 d), Na^{22} (2.6 y), and Pb^{210} (22.3 y). These isotopes are produced at different rates in different regions of the atmosphere, depending on altitude and geographic latitude. These characteristics of the production make it possible to use these isotopes for measuring the time scales of transport of air from one region of the atmosphere to another. The problem is not quite straightforward because there yet exist some uncertainties in measurement of the activities and ambiguities in the interpretation of a limited set of observations. The main uncertainty in the measurements arises from the fact that the volumes of air sampled are not accurately known; the error may be as large as 20 to 30 per cent (E. A. Ramo, private communication, 1960).² The uncertainties in the absolute concentrations of the activities in air are, therefore, uncertain by this amount. While interpreting the results one must consider the continuous (and possibly variable) production of the activities during the

past trajectories of the components of the air parcel which has been sampled. It is often difficult to do this unambiguously. However, we thought it worth while to investigate whether, in some simple situations, useful information could be derived from a reasonable number of measurements. As an attempt in this direction we measured the concentrations of Be^7 , P^{32} , and P^{33} at a few selected points in the atmosphere. The concentrations of another natural radioisotope, Pb^{210} (22.3 y), which arises from decay of radon (3.8 d) in the atmosphere, were also measured. Similar measurements were recently attempted as part of the *High Altitude Sampling Program* [1960] and by *List and Telegadas* [1961]. The observations, though not extensive, do indicate that the method is capable of yielding information essentially similar to that obtained from studies of nuclear bomb activities.

Experimental procedure. Air filters of type IPC-1478 through which large amounts of air (3×10^6 to 3×10^7 liters NTP) were filtered in some aircraft flights were supplied by E. A. Martell, Air Force Cambridge Research Center.

After adding the stable carriers of beryllium, phosphorus, lead, and cesium, the organic matter of the filters was destroyed by treatment with nitric and perchloric acids. Various activities were then isolated by using specific chemical procedures [Lal, Arnold, and Honda, 1960; Honda, Shedlovsky, and Arnold, 1961; Rama, Koide, and Goldberg, 1961]. The activities were measured by the usual methods of low-level counting. The purity of P^{32} and P^{33} sources was checked by measuring the decay of their activi-

¹ Leave from the Tata Institute of Fundamental Research, Bombay, India.
² According to a recent report [*High Altitude Sampling Program DASA 539*, 1961], the error has been reduced by recomputation of the flow rates in the filter papers. At the same time this correction indicates that about +10–20% corrections should be made for the activities in 9 out of 21 samples which we analyzed (Table 1). In this paper, however, all data are uncorrected.

TABLE 1. Observed Contents of Various Activities in the Samples

Date of Sampling	Location Lat., Long. (geog.), deg	Approx. Geomag. Lat., deg	Sampling Altitude, km	Tropopause Altitude, km	dpm/10 ⁴ liters (NTP)					
					Be ⁷	P ³²	Be ⁷ /P ³²	P ³² /P ³³	Pb ²¹⁰	Cs ¹³⁷
11/19/60	10N, 122E	0	4.6	16	0.46	0.02	22	0.67	0.01	
11/19/60	10N, 122E	0	7.6	16	.71	.03	25	0.50	.01	
11/19/60	10N, 122E	0	12	16	.81	.03	25	0.56	.004	
11/19/60	12N, 158W	10	15	16	11	.26	42	0.74	.06	
11/19/60	12N, 158W	10	21	16		.64		1.0	.11	44
1/26/61	29N, 98W	40	12	10	22	.53	41	0.80	.11	
1/26/61	29N, 98W	40	15	10	13	.42	30	0.6	.06	
1/26/61	29N, 98W	40	18	10	130	2.2	59	0.84	.21	35
1/26/61	29N, 98W	40	20	10	220	3.0	71	0.96		81
11/3/60	35N, 105W	45	4.6	11	2.0	.06	36	0.88	.07	
11/3/60	35N, 105W	45	12	11	27	.60	45	1.04	.13	
11/3/60	35N, 105W	45	15	12.5	39	.71	56	1.05	.14	3
11/3/60	35N, 105W	45	20	12.5	230	3.0	77	0.87	.07	96
11/2/60	42S, 147E	50	4.6	6.7	4.4	.08	54	0.91	.018	
11/2/60	42S, 147E	50	7.6	6.7	72	1.1	65	1.03	.13	6
11/2/60	42S, 147E	50	12	6.7	93	1.9	50	0.89	.13	12
11/2/60	42S, 147E	50	15	6.7	120	1.9	59	0.91	.13	22
11/2/60	42S, 147E	50	20	6.7	345	6.6	53	0.97	.14	190
2/27/61	64N, 149W	65	12	9.5	140	2.3	61	1.1	.15	
1/2/61	74N, 152W	70	3		4.1	.24	17	0.46	.13	
2/3/61	74N, 152W	70	3		2.2	.08	27	0.83	.16	
2/18/61	74N, 152W	70	3		4.7	.15	33	0.76	.17	
3/1/61	74N, 152W	70	3		2.6	.08	31	0.71	.19	

ties and the absorption of their β radiations. The decay was followed for a period of more than 2 months and was found to be characteristic for the two activities. The purity of the Be⁷ and Cs¹³⁷ sources was ascertained from their characteristic γ -ray spectra; that of Pb²¹⁰ sources was ascertained from the growth of Bi²¹⁰ activity in them.

Results. The results of various measurements are given in Table 1. The over-all experimental errors, excluding those due to uncertainty in the measurement of air volume, are believed to be less than 10 per cent, except for P³² determination which has an appreciable error due to the soft β ray. The error would be about 20 per cent in usual samples, and in the tropospheric samples of 10°N the figures were uncertain by as much as 50 per cent. The values of latitude are correct to within a few degrees, and of altitude within a few thousand feet. The height of the tropopause at the time of flight is also listed.

Some additional data on the activities of Be⁷, P³², P³³, S³⁵, and Na²² were obtained in preliminary experiments conducted in 1959 and 1960.

In these experiments there might have been somewhat higher uncertainty in the determination of absolute concentrations of isotopes in air because the experimental techniques were still in the development stage. The determinations of Be⁷/P³² and P³²/P³³ ratios were, however, comparable in accuracy to those in the 1960-1961 experiment; the error in S³⁵/P³² ratios might be high (~ 50 per cent). These data are summarized in Table 2. Regarding the Na²² data there was a possibility that the Na²² might also be derived from nuclear bomb explosions. For the sake of future comparison, these data are also included in Table 2.

Discussion. The following features of the observations on Be⁷, P³², and P³³ may be noted from Table 1:

(1) The concentration of each isotope in air increases with altitude and also with latitude.

(2) The values of Be⁷/P³² ratios in the tropospheric samples range from 30 to 77, and those for P³²/P³³ ratios range from 0.6 to 1.1; the mean values are 56 and 0.9, respectively. Most values are close to the mean values which may be taken as fairly representative for all higher

TABLE 2. Ratios of Various Observed Activities in Several Groups of Samples

g	Location		No. of Samples	Ratios of Disintegration Rate			
	Alt., km	Lat. (Geog.)		Be ⁷ /P ³²	P ³³ /P ³²	S ³⁵ /P ³²	Na ²² /P ³²
61	Stratosphere						
	12-21	10°-65°	13	56	0.93		
	12-20	40°	6	44	0.73	1.1	0.02
	15-21	15°-70°	6	82	0.9	2.0	≤0.03
*	14-24	45°	4	100	0.78		
61	Troposphere						
	3-16	10°-75°	10	31	0.76		

loon flight samples.

the samples; the values considerably lower than the mean are observed in the lower-altitude samples.

The values of Be⁷/P³² ratios in the tropospheric samples range from 17 to 54; the mean is 31, which is significantly lower than the mean found for the stratospheric samples.

These observations have to be interpreted by taking in view the rates of production of the isotopes at various points in the atmosphere. These rates are given by *Lal and Peters* as a function of altitude and of geographic latitude. They are reproduced for P³²

in Table 3; those for P³³ and Be⁷ can be obtained by multiplying the P³² rates by 0.82 and 100, respectively. These values are probably accurate to about 25 per cent. The activities of P³² observed in various experiments are also listed in Table 3 for comparison.

If the atmosphere were static, the steady-state activities at any latitude and altitude would be in equilibrium with the local production. However, any rapid transport of air between regions of dissimilar production will change the distribution of the activities. On the other hand, if the transport occurs on a time

TABLE 3. Comparison of the Production Rate of P³² in Air and the Measured Activities in Samples

Geomag.):	Production Rate : Atoms of P ³² /min/10 ⁴ liters of Air								
	0°	10°	20°	30°	40°	45°	50°	65°	90°
m	0.46	0.60	0.9	1.4	1.9	2.5	3.0	8.5	9.2
	.71	.88	1.0	1.8	2.7	3.5	5.0	7.4	7.8
	1.1	1.1	1.3	2.3	3.0	4.2	4.6	5.7	5.7
		(.64)*			(3.0)	(3.0)	(3.4)	(3.5)	
m	1.0	1.1	1.3	2.1	2.9	3.9	4.4	5.1	5.1
			(1.3)	(2.3)	(1.8)				
					(2.2)				
	.9	.9	1.2	1.8	2.4	2.7	3.0	3.1	3.1
m		(0.2)			(1.4)	(.7)	(1.9)		
					(.4)				
	.6	.6	.7	1.3	1.6	1.8	1.9	1.9	1.9
	(.03)				(.5)	(.6)	(1.9)	(2.3)	
m	.2	.2	.2	.3	.5	.5	.5	.5	.5
	(.03)				(.5)		(1.1)		
	.1	.1	.1	.1	.1	.1	.1	.1	.1
	(.02)				(.1)	(.1)	(.1)	(.1)	

The values in parentheses are the measured activities in dpm/10⁴ liters of air.

scale which is much longer than the half-life of the isotope under consideration, the observed distribution should not differ significantly from the calculated equilibrium distribution. In fact, this appears to be the case in most regions of the stratosphere for P^{32} and P^{33} activities, signifying that any large-scale transport, organized or turbulent, between two regions of dissimilar production must occur on a time scale much longer than weeks. The undersaturations observed in the low stratospheric samples, however, do seem to indicate the effects of stratospheric and tropospheric mixing.

The measured values of Be^7 activities in the stratosphere are only about 60 per cent of the values calculated for static stratosphere. It is possible that this discrepancy is entirely due to errors in calculations and in the experiment; an underestimate of 10 to 15 per cent in the Be^7 activities in experimental measurements and an overestimate of 25 per cent in the calculated values cannot be ruled out. Therefore, Be^7 activities also may well be in equilibrium with the local production in most regions of the upper stratosphere. An alternative explanation can be that the stratospheric Be^7 , from an altitude of 20 km and at all latitudes, is transported to the troposphere in a matter of a few months. If this is the case, the downward transport should result in a considerable increase in the concentration of Be^7 in the troposphere at high latitudes; the increase at low latitudes, however, is expected to be slight and would not be easily detectable. For that reason, measurements of Be^7 and P^{33} concentrations in tropospheric air and in wet fallout at high latitudes are especially desirable; the few measurements made by us at $74^\circ N$ and at 3 km altitude do not give any indication of the occurrence of such downward transport.

The observed magnitudes of the activities in the troposphere are rather low, and their vertical distribution is fairly uniform. This indicates the importance of the processes of turbulence and wet precipitation in the troposphere. The combined effect of these processes is to remove the particulate material from the troposphere in a matter of weeks; the observed undersaturations in the tropospheric Be^7/P^{33} and P^{33}/P^{32} ratios are consistent with this hypothesis.

The source of Pb^{210} in the atmosphere is the

decay of radon which enters the atmosphere after escaping from the ground, at an average rate of about 40 atoms/min/cm² of land area [Israel, 1951]. The inventories of radon and Pb^{210} should, therefore, each be 40 dpm/cm² of land area. The entire inventory of radon, which presumably is not subject to washout by rainfall, lies in the atmosphere, whereas most of the inventory of Pb^{210} , which is removed from the atmosphere by rainfall in a period short compared with its half-life, lies on the earth's surface. The results of our measurements on Pb^{210} in the atmosphere are summarized as follows.

(1) The concentrations of Pb^{210} in the stratosphere are 0.1 to 0.2 dpm/10⁴ liters and are markedly uniform. Even at $40^\circ S$, where the land area is comparatively small, the stratospheric concentration of Pb^{210} is similar to that observed in the northern hemisphere.

(2) The inventory of Pb^{210} in the stratosphere is of the order of 0.001 dpm/cm², which is a very small fraction (~ 0.01 per cent) of the average inventory of radon in the atmosphere.

(3) The stratospheric concentrations are generally higher than or comparable to those in the troposphere.

The loss of Pb^{210} from the stratosphere occurs by the processes of decay and of transfer to the troposphere. The transfer probably takes place by the exchange of air across the tropopause and the same process is presumably responsible for bringing the balancing amount of radon and its decay products (of which only Pb^{210} is of importance) from the troposphere into the stratosphere.

The magnitude of stratospheric inventory of Pb^{210} depends upon the concentration of radon and Pb^{210} in the air near the tropopause and on the rate of exchange of air across the tropopause. The former depends not only on the rate of exhalation of radon from the earth's surface but also on various meteorological factors which affect the upward transport of radon and downward removal of Pb^{210} . It is important to measure the concentrations of radon also for assessing the extent of mixing between the stratosphere and the troposphere.

From the markedly uniform concentrations in the stratospheric samples we conclude that the holdup time of Pb^{210} in the stratosphere is long enough to permit the processes of mixing to

the concentration more or less independent of altitude and latitude.

It may be remarked that the concentrations observed by *Burton and Stewart* [1960] in the stratosphere over the United Kingdom are higher by a factor of more than 5 than those reported by us in most regions of the stratosphere. Such high concentrations appear to be due to some local causes.

Acknowledgments. We are greatly indebted to Dr. James R. Arnold for his keen interest and encouragement throughout this work. We express our assurance in thanking Dr. E. A. Martell for his continuous cooperation in this program. The assistance of Dr. D. Lal and Dr. N. Yamagata in the early phases of this study in our laboratory is fully acknowledged. Skillful assistance of Miss M. Settle and Mr. M. Anderson is deeply appreciated. This research was supported by a contract with the U. S. Air Force, Geophysical Research Directorate.

REFERENCES

W. M., and N. G. Stewart, Use of long-

lived natural radioactivity as an atmospheric tracer, *Nature*, **186**, 584-589, 1960.

High Altitude Sampling Program, Defense Atomic Support Agency, Washington, D. C., 1960; 1961. Honda, M., J. P. Shedlovsky, and J. R. Arnold, Radioactive species produced by cosmic rays in iron meteorites, *Geochim. et Cosmochim. Acta*, **22**, 123-154, 1961.

Israel, H., Radioactivity of the atmosphere, *Compendium of Meteorology*, American Meteorological Society, Boston, Mass., 155-161, 1951.

Lal, D., J. R. Arnold, and M. Honda, Cosmic ray production rates of Be^7 in oxygen, and P^{32} , P^{33} , S^{35} in argon at mountain altitudes, *Phys. Rev.*, **118**, 1626-1632, 1960.

Lal, D., and B. Peters, Cosmic ray produced isotopes and their application to problems in geophysics, *Progress in Cosmic Ray and Elementary Particles Physics*, vol. 6, North Holland Publishing Co., Amsterdam (in press), 1961.

List, R. J., and K. Telegadas, The pattern of global atmospheric radioactivity, *Quarterly Summary Report*, U. S. Weather Bureau, April, 1961.

Rama, M. Koide, and E. D. Goldberg, Lead-210 in natural waters, *Science*, **134**, 98-99, 1961.

(Manuscript received July 1, 1961.)

Ground-Conductivity Determinations at Low Radio Frequencies by an Analysis of the Sferic Signatures of Thunderstorms

J. R. JOHLER AND C. M. LILLEY

*Central Radio Propagation Laboratory, National Bureau of Standards
Boulder, Colorado*

Abstract. A technique is described for determining the conductivity of the ground at low frequencies with the aid of sferic pulses from thunderstorms. The paper is illustrated with an actual conductivity determination, and the detailed comparison of waveforms predicted theoretically with those observed experimentally indicates that an effective value of conductivity can be measured to a precision of one or two significant figures. The results of the analysis indicate the application of both experimental and analytic techniques to other propagation studies, such as the evaluation of the reflection and transmission properties of the lower ionosphere.

Introduction. The predictability of the ground-mode of propagation is, in large measure, dependent upon a knowledge of the conductivity and dielectric constant of the large volume of ground of concern in any particular transmission. Considerable interest has developed in the conductivity of the ground at low frequencies, especially as a result of the Loran-C radio navigation system. Indeed, it has been shown [Smith-Rose, 1960] that the predictability of the Loran-C system, which operates on the ground-mode of propagation to distances of 1000 statute miles from the transmitters, can be greatly improved if the ground conductivity is known. A corresponding improvement in the synchronization with which time can be synchronized to the surface of the earth by time-locking clocks in a system to a common frame of reference (cesium beam standard) to a precision of nominal 0.1 microsecond instead of the actual 1 microsecond already accomplished in the absence of effective ground-conductivity measurements is possible if such values are known to an accuracy of one or two significant figures. The ground conductivity was measured in some detail by Smith-Rose [1934], and it is interesting to note that the ground-conductivity values in his measurements are practically a constant at low frequencies, whereas the dielectric constant exhibits a wide variation. Indeed, the dielectric constant grows to quite large numbers at extra-low frequencies (ELF). Nevertheless, the ratio of the dielectric constant term

$$(\omega^2/c^2)\epsilon_2$$

to the conductivity term

$$\left| i \frac{\omega^2}{c^2} \left[\frac{\sigma \mu_0 c^2}{\omega} \right] \right|$$

where $f = \omega/2\pi$ is the frequency, cycles per second; σ is the conductivity of the ground, mhos/meter; c is the speed of light, meters/second; μ_0 is the permeability of space, $4\pi(10^{-7})$ henry/meter; and ϵ_2 is the dielectric constant of the ground relative to a vacuum, where the wave number squared, k_2^2 , for waves in the ground,

$$k_2^2 = \frac{\omega^2}{c^2} \left[\epsilon_2 + i \frac{\sigma \mu_0 c^2}{\omega} \right]$$

which governs the propagation of the field, E , propagated a distance D meters, in the ground medium,

$$E = |E| \exp [-i\omega t + ik_2 D]$$

is quite small,

$$\epsilon_2 \omega / \sigma \mu_0 c^2 \ll 1$$

The work of Smith-Rose has recently been extended in great detail [Wait, 1959] at LF, VLF, ELF, and ULF, and the frequency dependence noted by Smith-Rose has been exploited so that the rate of change of effective conductivities of various rock materials has been utilized as an indication of the type of local rock

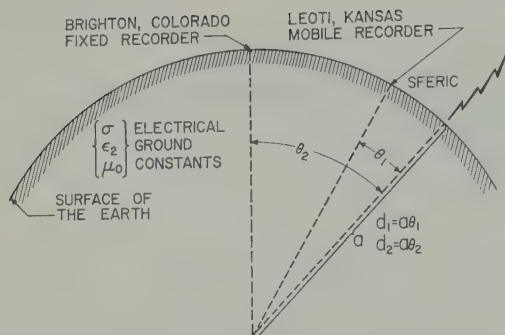


Fig. 1. Experimental arrangement for a ground-conductivity determination.

material. For example a very pronounced change in the frequency dependence was noted in the presence of sulfide mineralization in the ground. As a result, effective dielectric constants and effective conductivities were ascribed to various rocks with the implication that effective values correspond to various types of rock in the locality of the measurement. The measurement method proposed by *Wait and Conda* [1958] employed ground probes or electrodes inserted into drill holes in the ground. *Wait* [1957; see also *Wait and Campbell*, 1953] reports that dielectric constants of the order of 10^3 at 15 kc/s, or $\epsilon_2\omega/\sigma\mu_0c^2 = 0.025$, for example, are not uncommon. The role of the dielectric constant is certainly not negligible at low frequencies, especially if the conductivity or the effective conductivity is to be determined to more than one or two significant figures. In certain instances, especially in the Arctic, the dielectric constant can profoundly affect the ground-wave propagation. The analysis demonstrated in this paper can treat both the dielectric constant and the conductivity, although the displacement currents in the earth were so small in the examples illustrated that the value $\epsilon_2 = 15$ was an adequate assumption.

In the past, other methods for the measurement of ground conductivity or at least effective values of ground conductivity have been proposed. *Norton* [1937] has suggested the use of the wave-tilt method. This method employs a measurement of the ratio of the vertical electric field, E_v , and the horizontal field in the direction of propagation, of a continuous-wave (cw) signal, E_θ ; i.e., the wave tilt β is defined,

$$\tan \beta = E_\theta/E_v$$

which is proportional to conductivity, σ . Also the ground conductivity can be deduced by direct comparison of measured and theoretical values of amplitude and/or phase of propagating cw signals.

The effective ground conductivity of the continental United States has been investigated with the aid of Norton's methods [*Kirby*, 1951] and ground-conductivity maps [*Fine*, 1954] have been published. Such maps are deduced from measurements of the field strength of broadcast transmitters, and the applicability to low frequencies is questionable in view of the frequency dependence noted by *Smith-Rose* for medium and high frequencies.

The electromagnetic pulse radiation or 'sferic signatures' from thunderstorms or 'storm cells' has long been known to exist as a source of interference to radio communication, and extensive research has been in progress to evaluate such interference. Recently, however, the nature of the electromagnetic pulse waveform or 'sferic signature' and the nature of the source have been the subject of extensive study. In particular, the Eph system [*Hefley, Linfield, and Davis*, 1961] has provided advanced instrumentation for the location and identification of the source of the electromagnetic radiation from various parts of storms. The spectrum of frequencies investigated by *Hefley* and co-workers, which also corresponds to the frequency spectrum investigated in this paper, was confined to low frequencies (<150 kc/s), but this restriction does not preclude parallel medium- and high-frequency investigations.

Thus, it has been found to be quite possible to locate a position of a source and simultaneously record the signature of the source at various geographic locations. Obviously, such a technique can provide valuable instrumentation for the study of LF-VLF propagation phenomena.

The abundance of sferics, especially during the summer months over continental United States, for example, suggests the use of such sferic signals with the aid of the Eph system [*Hefley, Linfield, and Davis*, 1961] to measure the effective ground conductivity from simultaneous observations of sferic signatures at different geographic locations. Obviously, the effective conductivities deduced in this paper involve the large volumes of ground of concern in radio navigation-system (or other system which utilize

round wave to great distances) transmiss-
 Thus, the local effective conductivity
 is described in the above-mentioned methods
 of secondary interest. The method proposed
 of checking as to validity of application and
 large measure sidesteps the detailed analysis
 of local measurements to deduce an effective
 for a particular transmission.

Experimental procedure. One sferic signature
 recorder was located at Brighton, Colorado;
 another, in a mobile unit near Leoti, Kansas,
 100 miles from Brighton. At Brighton, the Eph-
 eum was used to select only the sferics from
 the sector centered on Leoti. Weather Bureau
 observations, including radar tracking of thunder-
 storms, were used to determine the distance
 from the Leoti station to the storm cells and to
 verify that the storms were approximately on
 the extension of the line joining Brighton and
 Leoti. The storm cell employed for this study
 produced a great number of sferics, of course,
 but only one was necessary for the conductivity
 determination and a second served as a check.
 Two vertical receiving antennas were used, the
 local electric component of the field, E_r , was
 observed. The aspect of cloud-to-ground lightning
 strokes is not precisely vertical. Therefore,
 strokes that occur on the extension of the line
 connecting the two observation points must be
 selected to eliminate any errors that might be
 introduced to stroke aspect. The effective source of
 radiation can thus be considered a vertically
 polarized source, since both recorders (Fig. 1)
 were observing the same sferic aspect.

The waveforms were permanently recorded
 on film, and the desired waveforms and cali-
 bration data were converted to digital form
 with the aid of an electronic scaling device (Ben-
 son-Lehner analog digital converter, 'Boscar').
 The waveforms taken simultaneously at
 the two recording sites were identified with the aid
 of the National Bureau of Standards WWV time
 signal, which was simultaneously recorded on the film.
 Distances between 100 and 1000 statute miles
 were considered to be the most desirable ranges
 for the conductivity determinations; the higher
 values of conductivity (>0.01) can be more
 precisely determined with the greater distances.
 The simple arrangement (Fig. 1) was employed
 to demonstrate the measurement principle;
 obviously more complicated arrangements or
 systems could be devised. The distances between

the Leoti sferic and the source, d_1 , and the
 Brighton sferic and the source, d_2 , were deter-
 mined from weather charts, wind-velocity charts,
 and radar observations of the storm cells in the
 area. The locations were approximately verified
 visually at the Leoti recorder.

The equipment was calibrated, before the
 commencement and upon the completion of the
 experiment, by two methods: (1) simultaneous
 recording of the same sferic pulse with the mobile
 recorder located in close proximity to the
 Brighton recorder (~ 1 mile) and a corresponding
 analysis of the transforms to determine the ratio
 of the transfer characteristics, $f_{r,2}(\omega)/f_{r,1}(\omega)$,
 where $f_r(\omega)$ is the complex transfer characteristic
 of the equipment and the subscripts 1 and 2
 refer to the mobile receiver-recorder and the
 Brighton receiver-recorder, respectively; and
 (2) analysis of the transforms of a damped sine-
 wave pulse introduced into the antenna circuit
 in such a manner as to simulate the induced
 emf of the field, and a simultaneous recording
 of this pulse together with the pulse at the
 output of the equipment. Thus, writing \mathcal{F} for
 the Fourier transformation (from time domain
 to frequency domain), assuming that the equip-
 ment is operated as a linear amplitude device,

$$\mathcal{F}[F(t)] = f(\omega)$$

where $F(t)$ is the induced emf test pulse, and
 the equipment transfer characteristic

$$f_r(\omega) = f_{out}(\omega)/f_{in}(\omega)$$

from which again the ratio $f_{r,2}(\omega)/f_{r,1}(\omega)$ can
 be determined.

Theory of pulse propagation. The propagated
 pulse can be described as a transient field,
 $E(t', d)$ [Johler and Walters, 1959], where the
 local time, $t' = t - \eta_1 d/c$, is used instead of the
 universal (source) time t , in which $\eta_1 \sim 1$ is the
 index of refraction of air at the surface of the
 earth, d is the distance in meters, c is the speed
 of light, $c \sim 3(10^8)$ m/sec. The transient field
 $E(t', d)$, is related to the time harmonic waves,
 $E(\omega, d)$, or the source, $F_s(t)$, assuming a linear
 amplitude response of the medium of propaga-
 tion, by the Fourier transform-integral

$$E(t', d) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(i\omega t') E(\omega, d) f_r(\omega) \\ \cdot \int_0^{\infty} \exp(-i\omega t) F_s(t) dt d\omega \quad (1)$$

in which the receiver-recorder mutilation function, $f_r(\omega)$, has been introduced to describe the action of such equipment on the form or shape of the pulse. In essence, the determination of the complex transfer characteristic,

$$f_r(\omega) = |f_r(\omega)| \exp [i \arg f_r(\omega)] \quad (2)$$

and its introduction into the operational procedure (1), comprises the equipment calibration, and the Fourier transform $f_x(\omega, d)$ can be written as the product of three factors,

$$f_x(\omega, d) = E(\omega, d) f_s(\omega) f_r(\omega) \quad (3)$$

in which the source transform $f_s(\omega)$ results from the first integration (1) and $E(\omega, d)$ is the transfer characteristic of the propagation medium, i.e., the more conventional solution for the field described by Maxwell's equations for a continuous time harmonic wave.

The source function $F_s(t)$ can be specified quite closely, and the corresponding fields $E(t', d)$ can be accordingly predicted. However, the source function $F_s(t)$ (ampere-meters) is not ordinarily observed or measured. It is more practical to measure the field, $E(t', d)$, at some short distance, d_1 , Figure 1. The waveform or sferic signature thus observed is, mathematically speaking, the real part of the amplitude-time function, $\text{Re } E(t', d)$.

Consider an experiment in which the signal, $\text{Re } E(t', d)$, is observed and recorded at some distance, d_1 , Figure 1, from the source. The theory is then required to predict the form of the signal recorded at some other distance, d_2 , Figure 1. The theory is also required to determine the form of the source, $F_s(t)$. The spectrum, $f_x(\omega, d_1)$ (3), can be determined directly from the observed signal, $\text{Re } E(t', d)$:

$$f_x(\omega, d_1) = \int_0^\infty \exp(-i\omega t') \text{Re } E(t', d_1) dt' \quad (4)$$

or

$$f_x(\omega, d) = \mathcal{F} \text{Re } E(t', d) \quad (5)$$

where \mathcal{F} designates the Fourier transform operation, Figure 2. The infinite integral can be split into the sum of several finite integrals, the intervals of integration being somewhat arbitrarily chosen but consistent with computation efficiency,

$$\begin{aligned} f_x(\omega, d_1) &= \int_0^{t'_{n_1}} F(\omega, t') dt' \\ &+ \int_{t'_{n_1}}^{t'_{n_2}} F(\omega, t') dt' + \cdots \\ &+ \int_{t'_{n_n}}^{t'_{n+1}} F(\omega, t') dt' + \cdots \end{aligned} \quad (6)$$

where each integral can be evaluated with gaussian quadrature,

$$\begin{aligned} &\int_{t'_{n_n}}^{t'_{n+1}} F(\omega, t') dt' \\ &= \sum_{m=1}^M W_m F(\omega, t'_m) + \epsilon(M) \end{aligned} \quad (7)$$

$m = 1, 2, 3, \cdots, M$, where the error term $\epsilon(M)$ can be made arbitrarily small by increasing M , and

$$t'_m = \frac{1}{2}(t'_{n+1} - t'_n)x_m + \frac{1}{2}(t'_{n+1} + t'_n) \quad (8)$$

$$W_m = \frac{1}{2}(t'_{n+1} - t'_n)H_m \quad (9)$$

The x_m 's are the gaussian abscissas, and M determines the number of values of the integrand to be used in the quadrature. The gaussian abscissas can be determined as the roots, x_m , of

$$\frac{d^M}{dx^M} (x^2 - 1)^M = 2^M M! P_M(x) \quad (10)$$

and the weights can then be determined,

$$H_m = \frac{2}{(1 - x_m^2)[P'_m(x_m)]^2} \quad (11)$$

where $P_m(x)$ is the Legendre function,

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$$

$$P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$$

$$P_4(x) = \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}$$

\vdots

(12)

Convenient tables of the abscissas x_m and the weights H_m have been developed [Davis and Rabinowitz, 1956].

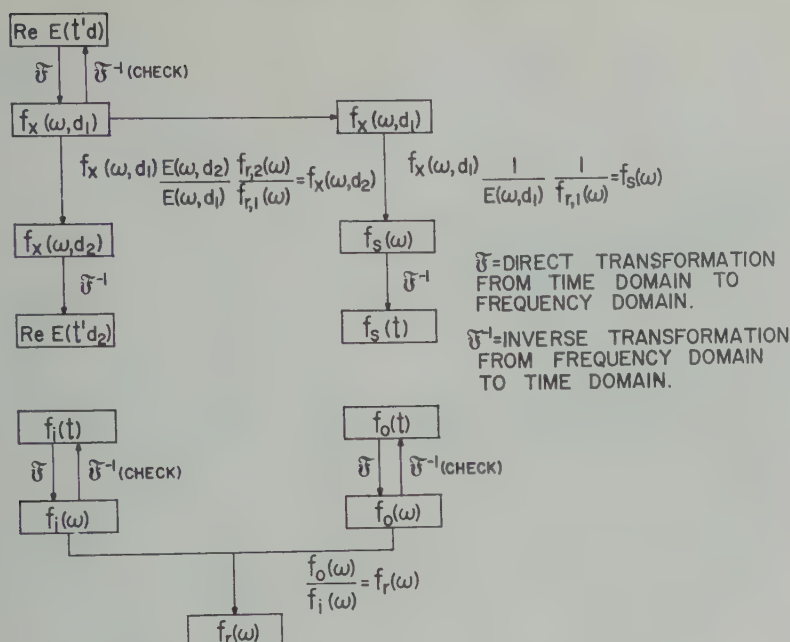


Fig. 2. Transformation flow chart for sferic signature analysis.

the spectrum of the source, $f_{x,s}(\omega)$, can then be determined with the aid of the transfer function of the medium, $E(\omega, d)$,

$$f_{x,s}(\omega) = \frac{f_x(\omega, d_1)}{E(\omega, d_1)} \frac{1}{f_{r,1}(\omega)} \quad (13)$$

the complex transfer function, $f_r(\omega)$, has been introduced to take account of the mutilation of the integrand or transform by the receiver-recorder equipment.

Since the real part of the signal, $\text{Re } E(t', d)$, is employed in the analysis, the source function $\text{Re } F(t)$, can be described as an integral with a symmetrical integrand,

$$F(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(i\omega t) f_{x,s}(\omega) d\omega \quad (14)$$

$$= \frac{1}{\pi} \int_0^{\infty} |f_{x,s}(\omega)| \cdot \{\cos[\omega t + \arg f_{x,s}(\omega)]\} d\omega \quad (15)$$

Therefore, (6), the infinite integral, can also be evaluated in the frequency domain somewhat arbitrarily but consistently with computation efficiency, replacing the integration limits with

$0 - \omega_1, \omega_1 - \omega_2 \dots$, so that the variable of integration becomes ω instead of t , and each integral is evaluated by a gaussian quadrature (7), where

$$\omega_m = \frac{1}{2}(\omega_{n+1} - \omega_n)x_m + \frac{1}{2}(\omega_{n+1} + \omega_n) \quad (16)$$

$$W_m = \frac{1}{2}(\omega_{n+1} - \omega_n)H_m \quad (17)$$

and the gaussian abscissas x_m and the weights H_m have previously been described (10)-(12).

The sferic signature, $E(t', d)$ (or any pulse, $\text{Re } E(t', d)$, for that matter), can then be predicted at the distance $d = d_2$ from the source, Figure 1, by a mutilation of the transform with the appropriate propagation medium, $E(\omega, d)$, and receiver transfer functions, $f_r(\omega)$,

$$f_x(\omega, d_2) = f_x(\omega, d_1) \frac{E(\omega, d_2)}{E(\omega, d_1)} \frac{f_{r,2}(\omega)}{f_{r,1}(\omega)} \quad (18)$$

where $f_{r,1}(\omega)$ and $f_{r,2}(\omega)$ are the transfer functions of the receiver-recorders at distances d_1 and d_2 respectively, Figure 1, and an inverse transformation,

$$E(t', d_2) = \frac{1}{\pi} \int_0^{\infty} |f_x(\omega, d_2)| \cdot \cos[\omega t' + \arg f_x(\omega, d_2)] d\omega \quad (19)$$

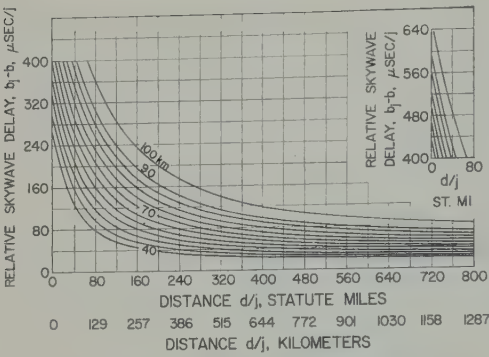


Fig. 3. Relative ionospheric wave delay, $b_i - b_0$, illustrating $j = 1$, the earliest local time, t_{j1} , on the pulse at which the time domain can be 'contaminated' with ionospheric waves.

or designating the operation of the inverse Fourier transformation or Fourier integral, \mathcal{F}^{-1} , Figure 2,

$$E(t', d_2) = \mathcal{F}^{-1}\{f_x(\omega, d_2)\} \quad (20)$$

which integral can be evaluated by the previously described quadrature procedure, (16)–(19), except that here the integrand becomes $F_t(\omega, d_2)$ instead of $F_t(\omega)$.

The recovery of the precise form of the source, $F_s(t)$, (15), is not always practical for d_1 quite large, since the ground wave is rather severely attenuated at high frequencies and the source is further obscured by the receiver-recorder equipment. The limit of resolution for which the source can be recovered at a given distance is experimental. But this merely means that the source is, as might be expected, obscured by the

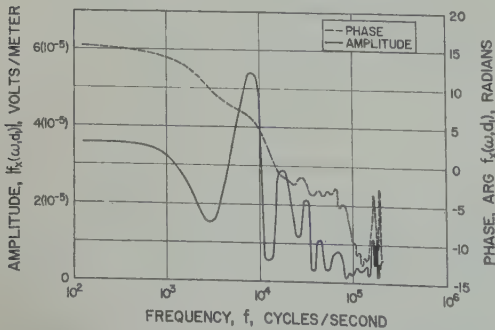


Fig. 4. Complex transform, $f_x(\omega, d_1)$ (amplitude and phase), of observed pulse, Leoti-1, illustrating the Fourier spectrum waves.

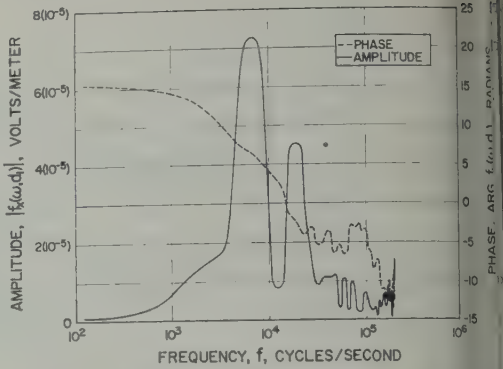


Fig. 5. Complex transform, $f_x(\omega, d_1)$ (amplitude and phase), of observed pulse, Leiti-2, illustrating Fourier spectrum waves.

propagation medium, $E(\omega, d)$, and the receiving equipment, $f_r(\omega)$. The exact form of the source is immaterial to this analysis, Figure 2, since prime interest is centered upon the predicted field, $E(\omega, d_2)$. Indeed, since the aspect of lightning current sources is rarely vertical, and the instrumentation employed vertical polarization, Figure 1, an effective vertically polarized component is alone of concern in this experiment. Thus, the sferics selected for this experiment were located on the extension of the line joining the two observation points, d_1, d_2 , Figure 1, presenting the same aspect or effective source to each receiver.

The ground-wave mode of propagation neglects the influence of the ionosphere. Indeed, the separation of the ground and ionospheric modes of propagation becomes increasingly difficult at

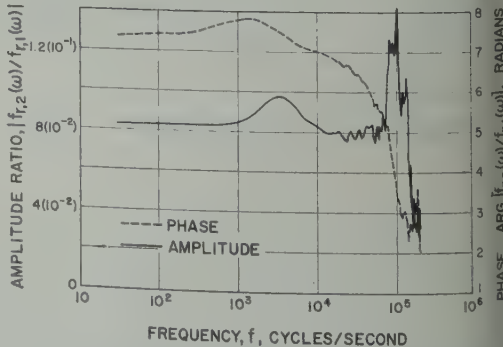
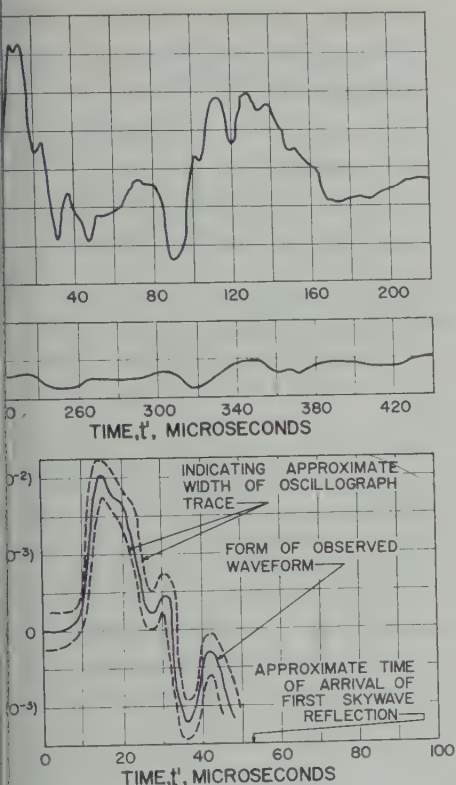


Fig. 6. Ratio of complex transfer characteristics, $f_{r,2}(\omega)/f_{r,1}(\omega)$ (amplitude and phase), of the recording equipment, illustrating the equipment calibration.



7. Observed spheric signatures distances d_2 for Leoti-1 pulse.

distances (>1000 miles) or lower frequencies (<10 kc/s). However, since the method resolves pulses in the time domain by use of the Fourier transformation, (19), (20), and Figure 2, since the observed pulse at a distance d_1 is very close to the source that a complete resolution of ground and sky-wave pulses is, the matter is resolved by a comparison of observed and predicted pulses, $E(t', d)$, in the time domain *before* the arrival of the sky-wave pulse.

As a consequence of ionospheric reflections, the modes of propagation, $j = 1, 2, 3$, are resolved upon the ground-wave time mode, $j = 0$. The composite pulse, $E(t', d) = E(t', d)$, assuming linear amplitude propagation is,

$$E(t') = \sum_{j=0}^p E_j(t', d) = \sum_{j=0}^p \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(i\omega t'_j) E_j(\omega, d) f_r(\omega) f_s(\omega) d\omega \quad (21)$$

where $j = 0, 1, 2, 3, \dots, p$, indicates the sum of a finite number of time modes which merely represents the sum of separate Fourier integrals for each time mode, separated in time by the sky-wave time-mode delay, t'_j . The earliest signal to arrive at the receiver is the ground wave, $E_0(t', d)$, and the first precursor of the ground-wave pulse can commence no sooner than a local time $t'_0 = 0$, where $t'_0 = t - b_0$, and $b_0 = \eta_1 d/c$. The higher-order time modes will arrive at even later times, t'_j ($j = 1, 2, 3, \dots$), where $t'_j = t - b_j$, and $b_j = \eta_1 D_j/c$ ($j = 1, 2, 3, \dots$), and D_j is some geometrical optical 'ray' length which is always greater than the distance d along the surface of the earth for ionospheric propagation. The quantity $b_j - b_0$, and especially $b_1 - b_0$, is called the relative sky-wave delay, Figure 3 (relative to the ground wave), and the quantity $b_1 - b_0$ at distances less than 1000 statute miles is the first time-mode sky-wave delay. This quantity represents the amount of time that can be utilized on the pulse $E(t', d_2)$ while comparing predicted ground-wave pulse parametric in conductivity with the observed pulse. The quantity $b_j - b_0$ is, of course, parametric in the reflection height, h , of the ionosphere, and at low frequencies a value of 68 km typifies daytime conditions and 80 km typifies nighttime conditions [Johler, Walters, and Lilley, 1960].

The ground-wave propagation medium, $E(\omega, d)$, can be calculated with the Watson, van der Pol, Bremmer series of residues, and the computation procedure employed in this paper has been detailed [Johler, Kellar, and Walters, 1956;

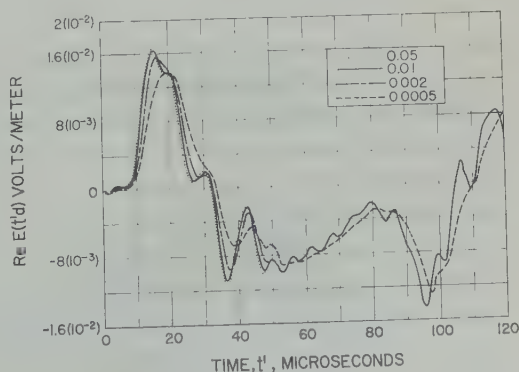


Fig. 8. Theoretical prediction of waveforms at distance d_2 , parametric in conductivity, σ , for observed pulse, Leoti-1, at distance d_1 .

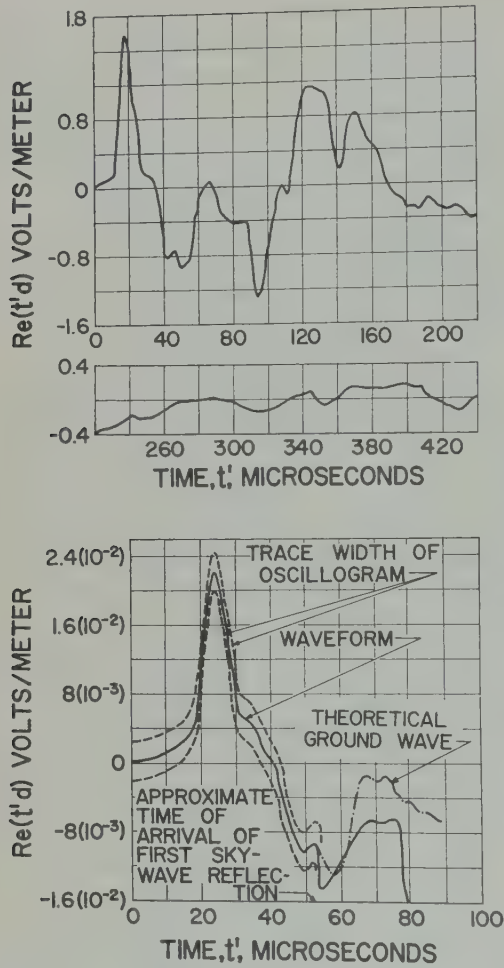


Fig. 9. Observed sferic signatures at distances d_1 and d_2 for Leoti-2 pulse.

Johler, Walters, and Lilley, 1959; Johler, Walters, and Lilley, 1960] designating the complex conjugate of $E(\omega, d)$, $E^*(\omega, d)$ (since these publications employed the time functions $\exp[-i\omega t]$),

$$E^*(\omega, d) = -i\omega C \left[2\pi\alpha^{2/3}(k_1a)^{1/3} \frac{d}{a} \right]^{1/2} \sum_{s=0}^{\infty} \frac{\exp \left\{ i \left[(k_1a)^{1/3} \tau_s \alpha^{2/3} \frac{d}{a} + \frac{\alpha d}{2a} + \frac{\pi}{4} \right] \right\}}{\left[2\tau_s - \frac{1}{\delta_s^2} \right]}$$

$s = 0, 1, 2, 3 \dots$, in which a is the radius of

the earth, $a \sim 6.367(10^6)$ meters, and

$$k_1 = (\omega/c)\eta_1 \tag{2'}$$

$$k_2^2 = \frac{\omega^2}{c^2} \left[\epsilon_2 + i \frac{\sigma\mu_0 c^2}{\omega} \right] \tag{2''}$$

$$\delta_s = \frac{i(k_2^2/k_1^2)\alpha^{1/3}}{(k_1a)^{1/3}[(k_2^2/k_1^2) - 1]^{1/2}} \tag{2'''}$$

$\mu_0 = 4\pi(10^{-7})$ henry/meter, $\alpha \sim 0.75 - 0.8$ and τ_s comprise the special roots of Riccati differential equation,

$$\frac{d\delta_s}{d\tau_s} - 2\delta_s^2\tau_s + 1 = 0 \tag{2iv}$$

which have been tabulated as a function of frequency [Johler, Walters, and Lilley, 1959]. Hence, the parameter τ_s , and in turn δ_s and k_2 introduce the effect of the dielectric constant ϵ_2 and the conductivity σ on the pulse. The ground wave has been tabulated at gaussian frequencies, $f_m = \omega_m/2\pi$, for various distances and conductivities [Johler, Walters, and Lilley, 1960]. For detailed conductivity and pulse propagation studies the entire analysis technique described above, including a residue series summation (as many as 1500 terms in the series were often required, especially at low frequencies and short distances) for the ground wave, was programmed in this study for large-scale electronic computers (IBM-704, CDC-1604).

Analysis of sferic signatures. The waveforms recorded at the shorter distances, d_1, d_2

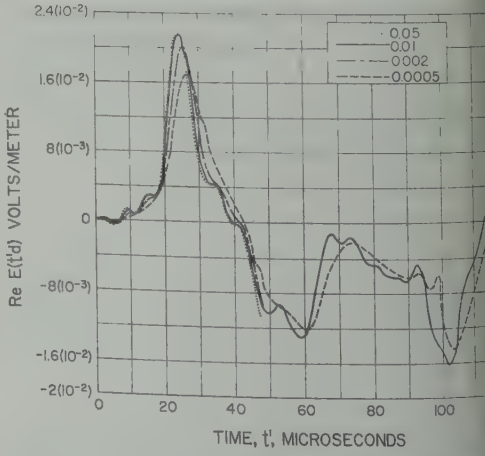


Fig. 10. Theoretical prediction of waveforms at distance d_2 , parametric in conductivity, σ , for observed pulse, Leoti-2, at distance d_1 .

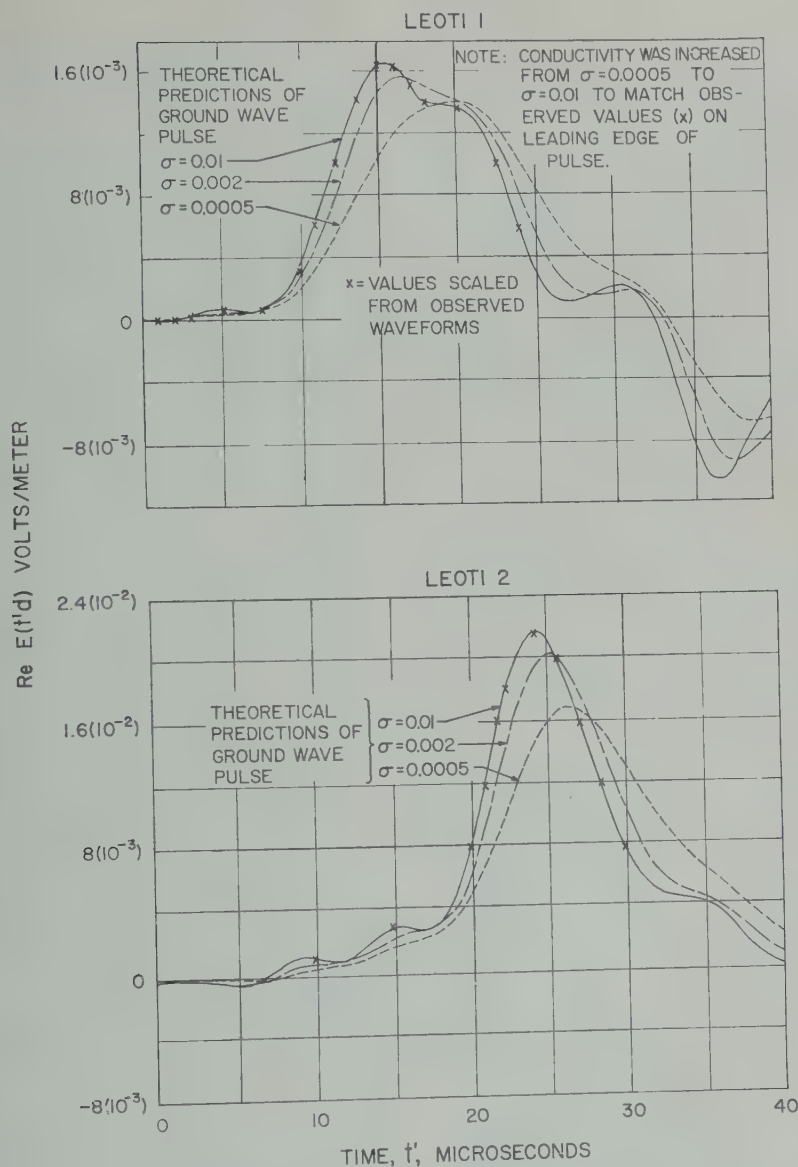


Fig. 11. Comparison of observed values along leading edge of the pulse with theoretical curves, illustrating the effective conductivity determination.

Figures 7 and 9, were analyzed by the direct, \mathcal{F} (Figs. 4 and 5), and inverse, \mathcal{F}^{-1} (Figs. 8, 10, 12, and 13), Fourier transformations described in Figure 2. The waveform was predicted at the distance d_2 (Figs. 7, 9, 11, 12, and 13), parabolic in conductivity σ . It was assumed that the displacement currents could be represented by the comparison of the leading edge of the

pulse waveform observed at distance d_2 was performed (Fig. 11) to estimate the conductivity by increasing the value of conductivity σ of the predicted pulse from 0.0005 to 0.01 mho/meter until the form of the predicted pulse matched the observed pulse.

An equipment calibration was introduced into the analysis, Figure 6, by forming the ratio of the complex transfer characteristics, Figure 2,

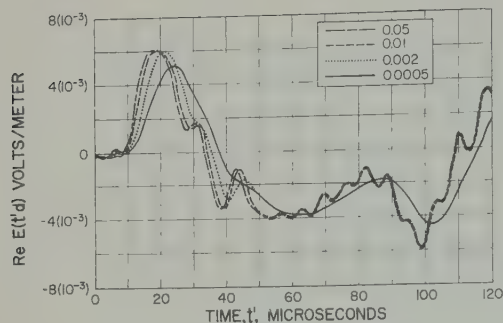


Fig. 12. Theoretical prediction of waveforms at a greater distance, $d_2 = 500$ statute miles, parametric in conductivity, σ , for observed pulse, Leoti-1, at distance d_1 .

$$f_{r,2}(\omega)/f_{r,1}(\omega)$$

from an analysis of sferic waveforms observed with both equipments recording together at the Brighton, Colorado, location the same sferic waveforms.

The complex transforms of the observed waveforms, $f_x(\omega, d_1)$, Figure 2, have been illustrated, Figures 4 and 5.

The comparison of the detail in the sferic, Figures 7, 9 and 11, observed at a distance, d_2 , with the predicted waveform, parametric in conductivity, led to the conclusion for both sferics that the conductivity of the western part of Kansas between Leoti, Kansas, and Brighton, Colorado, was approximately 0.01 or between 0.01 and 0.02 mho/meter. This seemed to be surprisingly high for land, which is typically 0.002 to 0.005. However, both sferic signatures gave the same value of conductivity, Figure 11. Since the conductivity was quite high, it would be desirable to employ a greater distance, d_2 , for its measurement, since the pulse stretching, Figures 12 and 13, would be increased by the greater distance, and the conductivity measurement would be more sensitive to the change in the pulse detail. It would also be quite desirable to improve the precision of the recorded oscillograms, since the trace, Figures 7 and 9, was quite wide for the observed pulses at the distance d_2 . It was for this reason that the pulses observed at the Leoti mobile recorder were employed in the analysis. Careful scaling of the detail of the wide trace at the distance d_2 revealed a most remarkable agreement between the predicted and observed pulses provided that

the equipment calibration, Figure 6, was employed in the analysis. A check of other independent measurements of conductivity in the area yielded the value $\sigma = 0.015$, which closely agrees with our conclusions.

The Leoti-1 pulse was re-evaluated by chopping the pulse off at 48 microseconds, a time early enough to eliminate all sky waves, Figure 3. This, however, introduced a step function into the transform, $f_x(\omega, d)$. But, upon performing the inverse transformation, the pulse again sort of out this fictitious phenomenon in the time domain, and again the same conductivity was deduced by matching the leading edge of the pulse. It was therefore concluded that the errors introduced by any sky-wave contamination which could be introduced into the transform was nil.

The quadrature techniques, which are quite sensitive to the changes in detail of the pulse, appear to be a delicate operation. Precision was assured, however, by various tests for convergence of the integrals (4) and (19), and any errors that could appear could only result from inaccuracies in the recording equipment and the electronic scaling procedure employed on the observed waveforms. In spite of the comparatively crude measurements made to demonstrate the principle, a precision of at least one significant figure was readily obtained. Further precision could be obtained by an analysis of a large number of waveforms with greater distances, d_2 , Figures 12 and 13, so that the pulse stretching is more obvious. More precise recording equipment would also be desirable.

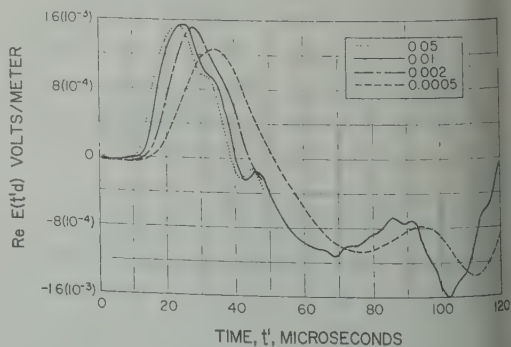


Fig. 13. Theoretical prediction of waveforms at a greater distance, $d_2 = 1000$ statute miles, parametric in conductivity, σ , for observed pulse, Leoti-1, at distance d_1 .

ussion. The experiment described indicates that an analysis of a large number of forms employing precision recordings and distances, d_2 , can be used to determine values of ground conductivity for large areas of ground with an accuracy of one or two significant figures. Since the earth is not homogeneous, the application of the theory of a homogeneous, smooth, spherical earth over a range of frequencies (10 to 200 kc/s) over inhomogeneous, irregular terrain is implied. Also, a large volume of experimental data has been accumulated over such terrain in eastern United States employing a pulsed radio navigation system [Hefley, 1960], and the behavior of the system was explained by ascribing an effective value of ground conductivity to the large areas of concern in the system. Greater precision for the system over land would, of course, require a detailed analysis of the irregularities and inhomogeneities in the terrain of concern in the propagation circuit, and indeed such an analysis is considered to be a very worthwhile experimental and theoretical effort. In the absence of such a detailed analysis of each propagation circuit, however, quite remarkable stability of the system timing (phase) was achieved for the pulsed system.

Similarly, in the case of spheric pulses, it is a necessary condition that the use of the effective value of conductivity in the theory of a homogeneous, smooth, spherical earth be capable of predicting the details of the pulse. All pulses analyzed to date, even over inhomogeneous terrain, have met this necessary condition. The employment of such analysis over a wide variety of cases with wide variations in conductivity and terrain will determine the limits under which the technique and notion of an effective value are sufficient.

Of course, if the necessary condition, that the pulse be faithfully predicted at a distance d_2 , is not satisfied to the required precision, the part of the transform $E(\omega, d)$, (1), must be replaced by a suitable transfer function for the heterogeneous earth or irregular terrain, and, assuming a correct effective conductivity and dielectric constant values or suitable dimension values for irregular terrain, the necessary condition that the pulse $E(t, d)$ be predicted must again be satisfied. The details for the construction of such transfer functions, $E(\omega, d)$, for hetero-

geneous, irregular terrain are beyond the scope of this paper.

The time on the waveform, greater than the sky-wave delay, Figure 3 ($b - b_1 \sim 55$ microseconds), at distances d_2 , is contaminated with the sky-wave reflection, Figures 7 and 9. The analysis described has determined the form of the ground-wave pulse. The ground-wave pulse can be removed from the analysis, since the shape of the pulse is known for all time, 0-120 microseconds, Figures 8, 10, 12, and 13. The remainder of the pulse comprises the first time mode, $j = 1$, which has been reflected from the ionosphere. A similar transient analysis, in which the transform for the sky wave (23), $E_j(\omega, d)$, $j = 1$, can be employed to deduce the reflection coefficient of the ionosphere, but this task is considered beyond the scope of this paper.

The sky-wave delay ($\sim 55-60$ microseconds), Figures 7 and 9, corresponds to a reflection height of only 55-60 km. Such a short sky-wave delay was especially surprising because the measurements were made at night (2100 MST, May 23, 1960). Possibly this phenomenon indicates a disturbed condition of the ionosphere. Such short sky-wave delay times, $b_1 - b_0$, were not observed on previous nights on several spheric signatures examined but not employed in this analysis.

The analysis could be employed to evaluate the dielectric constant. Thus, a family of curves similar to Figures 8, 10, 12, and 13 could be generated for study of any modification of pulse detail as a result of the dielectric constant. The dielectric constant was not considered to be important in this experiment, since a nominal value, $\epsilon_2 \sim 15$, was assumed, and since such a value would not affect the first or second significant figure in the conductivity determination.

Conclusions. A necessary condition for use of the theory of a smooth, homogeneous, spherical earth, with an effective value of conductivity and dielectric constant, to predict the behavior of a pulse transmission over inhomogeneous, irregular, terrain is that the detailed form or shape of the pulse be faithfully determined by the prediction at least to the recording accuracy of two significant figures. The limits for which such a notion is sufficient can be determined only after a wide variety of pulses propagated over a wide variety of terrains have been examined in detail by the techniques described.

The failure to satisfy the necessary condition requires the examination of each particular transmission in detail and substituting a transfer function, $E(\omega, d)$, for the ground wave over heterogeneous, irregular earth, into the transient solution. Again, the necessary condition that the pulse be faithfully reproduced in detail must be satisfied. Subject to these restrictions, a precision ground-conductivity determination is possible with these methods to a significance of one or two figures.

The technique for the analysis of transients applied in this paper can be employed in researches of the types, such as the deduction of the reflection coefficient of the ionosphere by a removal of the ground-wave pulse from the observed waveform at greater distances from the source of the sferic.

Acknowledgments. The field measurements were made by R. F. Linfield, C. A. Samson, E. L. Berger, G. F. Schreiber, E. E. Johnson, and M. W. Schroeder. Special acknowledgment is made of the assistance of C. A. Samson, R. F. Linfield, and G. F. Schreiber in analyzing the data.

REFERENCES

- Davis, P., and P. Rabinowitz, Abscissas and weights for Gaussian quadratures of high order, *J. Research NBS*, 56, 35-37, 1956.
- Fine, H., An effective ground conductivity map for continental United States, *Proc. IRE*, 42(9), 1405-1408, 1954.
- Hefley, G., Timing potentials of Loran-C, *Signal J. Armed Forces Commun. and Electronics Assoc.*, 15(6), 45-47, 1961. See also R. H. Doherty, G. Hefley, and R. F. Linfield, Timing potentials of Loran-C, *Proc. Fourteenth Annual Symposium on Frequency Control*, Atlantic City, N. J., sponsored by Frequency Control Division, U. S. Army Signal Research and Development Lab., Fort Monmouth, N. J., pp. 276-297, May 31-June 2, 1960.
- Hefley, G., R. F. Linfield, and T. L. Davis, The Eph system for VLF direction finding, *J. Research NBS*, 65D, 2 March-April 1961.
- Johler, J. R., On LF ionospheric phenomena in radio navigation systems, paper for Avionics Panel Meeting of the Advisory Group for Aeronautical Research and Development (AGARD), October 3-8, 1960, Istanbul, Turkey, of organization du Traité de l'Atlantique Nord, 64 rue de Varenne, Paris 7 (to be published, Pergamon Press).
- Johler, J. R., W. J. Kellar, and L. C. Walters, Phase of the low radiofrequency ground wave, *NBS Circ. 573*, U. S. Government Printing Office, Washington, D. C., June 26, 1956.
- Johler, J. R., and L. C. Walters, Propagation of the ground wave pulse around a finitely conducting spherical earth from a damped sinusoidal source current, *IRE Trans. on Antennas and Propagation*, AP-7, 1, 1-10, January 1959.
- Johler, J. R., L. C. Walters, and C. M. Lilley, Low- and very low-radiofrequency tables of ground wave parameters for the spherical earth theory: the roots of Riccati's differential equation, *NBS Tech. Note 7*, PB 151366, U. S. Dept. of Commerce, Office of Technical Services, Washington, D. C., February 1, 1959.
- Johler, J. R., L. C. Walters, and C. M. Lilley, Amplitude and phase of the low- and very low-radiofrequency ground wave, *NBS Tech. Note 60*, PB 161561, U. S. Dept. of Commerce, Office of Technical Services, Washington, D. C., June 1, 1960.
- Kirby, R. S., J. C. Harman, F. M. Capps, and R. N. Jones, Effective radio ground-conductivity measurements in the United States, *NBS Circ. 546*, U. S. Government Printing Office, Washington, D. C., February 26, 1954.
- Norton, K. A., The propagation of radio waves over the surface of the earth and in the upper atmosphere, part II, *Proc. IRE*, 25(2), 1203-1236, 1937.
- Smith-Rose, R. L., Electrical measurements on soil with alternating currents, *J. Inst. Elec. Engrs.*, 75, 221-237, 1934.
- Wait, J. R., The effective electrical constants of soil at low frequencies (letter to editor), *Proc. IRE*, 45, 10, 1957.
- Wait, J. R., editor, *Overvoltage Research and Geophysical Applications*, International Series of Monographs on Earth Sciences, 4, Pergamon Press, New York, 1959.
- Wait, J. R., and L. L. Campbell, Effect of a large dielectric constant on ground-wave propagation, *Can. J. Phys.*, 31, 456-457, 1953.
- Wait, J. R., and A. M. Conda, On the measurement of ground conductivity at VLF, *IRE Trans. on Antennas and Propagation*, AP-6, 3, 273-277, 1958.

(Manuscript received July 6, 1961.)

S_q and Ocean

TSUNEJI RIKITAKE¹

Geophysical Institute, Istanbul University, Istanbul, Turkey
and

Seismological Institute, Technical University of Istanbul, Istanbul, Turkey

Abstract. A theory of electromagnetic induction within a hemispherical conducting sheet over a nonconductor and underlain by a concentric sphere of uniform conductivity is described. The theory is applied to the induction by S_q in a vast ocean. It is concluded that the electric currents induced in the ocean are considerably smaller than those estimated for a single hemispherical sheet, so that the electromagnetic coupling between the ocean and the conducting part of the earth's mantle cannot be neglected in a study of this kind. The anomalous magnetic field that is caused by the ocean is so small that its maximum value hardly exceeds gammas. The ocean effects on S_q that have been studied so far on the basis of a single sheet would certainly be an overestimate.

Introduction. Considerable progress has been in recent years in theories on the possible influence of the ocean on transient geomagnetic effects. One of the most important conclusions is that the self-induction of a vast ocean plays a role so important that we can hardly estimate the electric currents induced in the ocean (Rikitake, 1960) unless appropriate account is taken of the self-induction.

At this point, also made clear by Rikitake [1961], the electric currents induced in a hemispherical ocean by a rapid geomagnetic change are greatly reduced by the presence of the conducting part of the earth's mantle, which seems to be even more important. It is certain that the effect of the ocean estimated in regard to a hemispherical conducting sheet has been an overestimate. Since the electromagnetic coupling between the ocean and mantle was derived only for the case of a very rapid change [Rikitake, 1961], it is desirable to examine the extent to which the magnetic daily variation, or S_q (one of the typical variations in the earth's magnetic field), is affected by an ocean of large scale. What the writer would like to present here is a study of electromagnetic induction in a hemispherical ocean by S_q , extended to a case in which suitable account is taken of the conducting part of the mantle.

Model of an ocean and an inducing field of S_q . We take a hemispherical conducting sheet at

on leave from the Earthquake Research Institute, Tokyo University, Tokyo, Japan.

the earth's surface as a model of a large ocean like the Pacific. We then assume that the part of the earth's mantle with radius smaller than qa (a is earth's radius and q is the ratio of the radius of the conducting mantle to a) has a uniform conductivity σ [Chapman, 1919; Chapman and Price, 1930; Lahiri and Price, 1939; Rikitake, 1950]. The system of conductors is shown in Figure 1.

As in a previous paper [Rikitake, 1960], the magnetic potential of the inducing field of S_q arising from outside the earth may be taken as

$$W_s = a \sum_n \sum_m (r/a)^n (e_{n,c}^m \cos m\alpha\tau + e_{n,s}^m \sin m\alpha\tau) P_n^m(\cos \theta_0) \quad (1)$$

where α , τ , and θ_0 denote, respectively, the angular measure of time, local time, and colatitude. The terms $e_{n,c}^m$ and $e_{n,s}^m$ can be computed from the results of a spherical harmonic analysis of S_q . P_n^m is Schmidt's spherical function.

If we denote by t the time at the $\phi_0 = 0$ meridian, on the assumption that the $\phi_0 = 0$ meridian agrees with the Greenwich one, we have

$$\alpha\tau = \alpha t + \phi_0 \quad (2)$$

so that (1) leads to

$$W_s = a \sum_n \sum_m (r/a)^n \{e_{n,c}^m \cos m(\alpha t + \phi_0) + e_{n,s}^m \sin m(\alpha t + \phi_0)\} P_n^m(\cos \theta_0) \quad (3)$$

which, by adopting a complex expression,

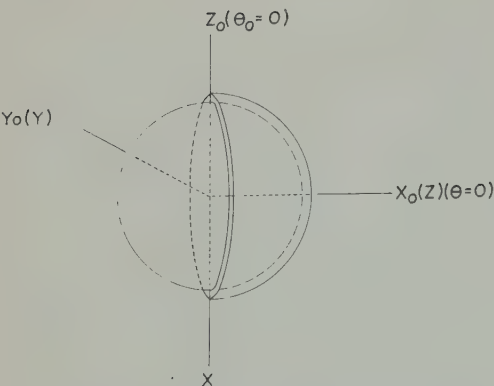


Fig. 1. Hemispherical ocean, conducting mantle, and coordinate systems.

becomes

$$W_e = a \sum_n \sum_m (r/a)^n (e_{n,c}^m - ie_{n,s}^m) \cdot e^{im(\alpha t + \phi_0)} P_n^m(\cos \theta_0) \tag{4}$$

Under the condition that suitable account is taken of the real and imaginary parts of the final results, (4) may be conveniently taken as the inducing field of S_a .

The coordinate system (x_0, y_0, z_0) is to be transformed to (x, y, z) as is also shown in Figure 1, in order to have solutions of good convergence. Accordingly, the ocean occupies the part of the spherical surface (radius = a) defined by $0 < \theta < \pi/2$. The relations between the rectangular and polar coordinates are taken as is customary.

With the aid of transformation formulas [Satō, 1950; Rikitake, 1960], we obtain

$$\left. \begin{aligned} P_2^1(\cos \theta_0) \cos \phi_0 &= -P_2^1(\cos \theta) \cos \phi \\ P_2^1(\cos \theta_0) \sin \phi_0 &= P_2^2(\cos \theta) \sin 2\phi \\ P_3^2(\cos \theta_0) \cos 2\phi_0 &= \frac{\sqrt{10}}{4} P_3^1(\cos \theta) \cdot \cos \phi + \frac{\sqrt{6}}{4} P_3^3(\cos \theta) \cos 3\phi \\ P_3^2(\cos \theta_0) \sin 2\phi_0 &= -P_3^2(\cos \theta) \sin 2\phi \end{aligned} \right\} \tag{5}$$

where θ_0 and ϕ_0 are the original coordinates.

We may take only the 24-hourly and 12-hourly components of S_a as the inducing field because they are most important for the mean S_a . In that case, we write

$$W_e = W_1 + W_2 \tag{6}$$

Applying the transformation formulas to (6) we obtain

$$\begin{aligned} W_1 &= a(r/a)^2 e^{i\alpha t} (e_{2,c}^1 - ie_{2,s}^1) [-P_2^1(\cos \theta) \cdot \cos \phi + iP_2^2(\cos \theta) \sin 2\phi] \\ W_2 &= a(r/a)^3 e^{2i\alpha t} (e_{3,c}^2 - ie_{3,s}^2) \cdot \left[\frac{\sqrt{10}}{4} P_3^1(\cos \theta) \cos \phi - iP_3^2(\cos \theta) \cdot \sin 2\phi + \frac{\sqrt{6}}{4} P_3^3(\cos \theta) \cos 3\phi \right] \end{aligned} \tag{7}$$

According to Chapman's analysis for the mean equinox in 1905 [Chapman, 1919], the coefficients are those as given in Table 1. The time origin is taken at local midnight.

Theory of electromagnetic induction. In a nonconducting region we may define a magnetic potential W which satisfies

$$\nabla^2 W = 0 \tag{8}$$

Typical terms of W for $r > a$ and for $a > r > q$ are, respectively, written as

$$W_n^m = a(e_n^m \rho^n + i_n^m \rho^{-n-1}) S_n^m \text{ for } \rho > 1 \tag{9}$$

and

$$W_n^m = a(e_n^{m'} \rho^n + i_n^{m'} \rho^{-n-1}) S_n^m \text{ for } 1 > \rho > q \tag{10}$$

where

$$\rho = r/a \tag{11}$$

and S_n^m stands for either $P_n^m(\cos \theta) \cos m\phi$ or $P_n^m(\cos \theta) \sin m\phi$. The three components of the magnetic field then become

$$\left. \begin{aligned} H_r &= -\{ne_n^m \rho^{n-1} - (n+1)i_n^m \rho^{-n-2}\} S_n^m \\ H_\theta &= -(e_n^m \rho^{n-1} + i_n^m \rho^{-n-2}) \partial S_n^m / \partial \theta \\ H_\phi &= -(e_n^m \rho^{n-1} + i_n^m \rho^{-n-2}) \partial S_n^m / (\sin \theta \partial \phi) \end{aligned} \right\} \text{ for } \rho > 1 \tag{12}$$

TABLE 1. The Coefficients for the Harmonics of the Magnetic Potential of S_a , in Gammas (after Chapman)

$e_{2,c}^1$	$e_{2,s}^1$	$e_{3,c}^2$	$e_{3,s}^2$
7.1	-3.1	-4.1	1.7

$$\left. \begin{aligned} & - \{ n e_n^{m'} \rho^{n-1} \\ & - (n+1) i_n^{m'} \rho^{-n-2} \} S_n^m \\ & - (e_n^{m'} \rho^{n-1} \\ & + i_n^{m'} \rho^{-n-2}) \partial S_n^m / \partial \theta \\ & - (e_n^{m'} \rho^{n-1} \\ & + i_n^{m'} \rho^{-n-2}) \partial S_n^m / (\sin \theta \partial \phi) \end{aligned} \right\} \quad \text{for } 1 > \rho > q \quad (13)$$

the conducting region ($r < qa$), we take vector potential \mathbf{A} which satisfies

$$\nabla^2 \mathbf{A} = 4\pi\sigma \partial \mathbf{A} / \partial t \quad (14)$$

the magnetic permeability is assumed as in electromagnetic units. A typical solution is given as

$$\mathbf{A}_n^m = a f_n(t, \rho) (\mathbf{r} \times \text{grad } S_n^m) \quad (15)$$

which the components of the magnetic are also calculated as

$$\left. \begin{aligned} & - \rho^{-1} n(n+1) f_n S_n^m \\ & - \rho^{-1} \partial(\rho f_n) / \partial \rho \partial S_n^m / \partial \theta \\ & - \rho^{-1} \partial(\rho f_n) / \partial \rho \partial S_n^m / (\sin \theta \partial \phi) \end{aligned} \right\} \quad \text{for } q > \rho \quad (16)$$

function f_n satisfies the differential equation

$$(f_n / d\rho) / d\rho = \{ n(n+1) + k^2 a^2 \rho^2 \} f_n \quad (17)$$

$$k^2 = 4\pi\sigma p \quad (18)$$

where p is the time-derivative operator $\partial/\partial t$. In the case of a uniformly conducting sphere, it is known that f_n is given as [Chapman and Bartels, 1940]

$$f_n = C_n \rho^n F_n(k^2 \rho^2 a^2) \quad (19)$$

where C_n is a constant and $F_n(k^2 \rho^2 a^2)$ is essentially a modified Bessel function. F_n has been studied in detail by Chapman and Bartels [1940]. When we assume that the current function in the current sheet at $r = a$ is expressed as

$$\Psi = \sum_n \sum_m K_n^m S_n^m \quad (20)$$

the continuity condition of the magnetic field at $r = a$ gives

$$\left. \begin{aligned} n e_n^m - (n+1) i_n^m &= n e_n^{m'} - (n+1) i_n^{m'} \\ e_n^m + i_n^m &= (4\pi/a) K_n^m + e_n^{m'} + i_n^{m'} \end{aligned} \right\} \quad (21)$$

From the similar condition at $r = qa$ we also obtain

$$\left. \begin{aligned} n e_n^{m'} - (n+1) q^{-2n-1} i_n^{m'} \\ &= n(n+1) C_n F_n(k^2 q^2 a^2) \\ e_n^{m'} + q^{-2n-1} i_n^{m'} &= C_n [(2n+1) \\ &\cdot F_{n-1}(k^2 q^2 a^2) - n F_n(k^2 q^2 a^2)] \end{aligned} \right\} \quad (22)$$

Solving (22) we obtain

$$e_n^{m'} = (n+1) C_n F_{n-1}(k^2 q^2 a^2) \quad (23)$$

and

$$i_n^{m'} = n q^{2n+1} C_n \cdot [F_{n-1}(k^2 q^2 a^2) - F_n(k^2 q^2 a^2)] \quad (24)$$

Putting (23) and (24) into (21) and solving it with respect to e_n^m and i_n^m , we obtain

$$\left. \begin{aligned} e_n^m &= \frac{n+1}{2n+1} \frac{4\pi}{a} K_n^m \\ &+ (n+1) C_n F_{n-1}(k^2 q^2 a^2) \end{aligned} \right\} \quad (25)$$

and

$$\left. \begin{aligned} i_n^m &= \frac{n}{2n+1} \frac{4\pi}{a} K_n^m + n q^{2n+1} C_n \\ &\cdot [F_{n-1}(k^2 q^2 a^2) - F_n(k^2 q^2 a^2)] \end{aligned} \right\} \quad (26)$$

Since e_n^m is given as the inducing field, (25) specifies the relation between K_n^m and C_n . If K_n^m is obtained somehow, it is possible to estimate i_n^m in terms of e_n^m .

We are now in a position to determine K_n^m . According to the theory of electromagnetic induction within a thin sheet [Price, 1949; Rikitake and Yokoyama, 1955], the condition that should be satisfied by the current function on the sheet is

$$\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin \theta} \frac{\partial^2}{\partial \phi^2} \right] \Psi = a^2 K \frac{\partial H_r}{\partial t} \quad (27)$$

TABLE 2. The Coefficients for the Current Functions, in Gammas, Multiplied by $a/4\pi$ in the Hemispherical Ocean in the Presence of the Conducting Mantle

n	\bar{a}_n	a_n^*	\bar{b}_n	b_n^*	$\bar{\alpha}_n$	α_n^*	$\bar{\beta}_n$	β_n^*	$\bar{\gamma}_n$	γ_n^*
1	0.158	-1.911			-0.073	0.037				
2	0.085	-0.958	-0.976	-0.222	-0.038	-0.713	-1.429	0.079		
3	0.029	-0.298	-0.478	-0.044	-0.021	-0.820	-1.026	0.049	-0.007	-0.64
4	0.001	-0.001	0.000	0.018	-0.016	-0.402	-0.378	0.014	-0.014	-0.33
5	-0.004	0.037	0.098	0.022	-0.008	0.003	-0.001	0.000	-0.009	0.00
6	0.000	0.000	0.000	0.000	0.002	0.091	0.057	-0.001	-0.004	0.08

where K and H_r are the total conductivity of the sheet and the component of magnetic field normal to the surface, respectively, so that on the conducting part of surface $r = a$ we have

$$\begin{aligned} \sum_n n(n+1)K_n^m S_n^m \\ = a^2 K p \sum_n [n e_n^m - (n+1) i_n^m] S_n^m \\ = a^2 K p \sum_n n(n+1) C_n \{ F_{n-1}(k^2 q^2 a^2) \\ - q^{2n+1} [F_{n-1}(k^2 q^2 a^2) \\ - F_n(k^2 q^2 a^2)] \} S_n^m \end{aligned} \quad (28)$$

Eliminating C_n from (28) with the aid of (25), we obtain

$$\begin{aligned} - \sum_n n(n+1)K_n^m S_n^m = a^2 K p \sum_n n \\ \cdot \left(\frac{n+1}{2n+1} \frac{4\pi}{a} K_n^m - e_n^m \right) \{ 1 - q^{2n+1} \\ \cdot [1 - F_n(k^2 q^2 a^2) / F_{n-1}(k^2 q^2 a^2)] \} S_n^m \end{aligned} \quad (29)$$

Attention should be paid to the fact that (29) is true only for $0 < \theta < \pi/2$ at $r = a$, whereas

$$\sum_n n(n+1)K_n^m S_n^m = 0 \quad (30)$$

holds good for both halves of the surface or $\pi/2 < \theta < \pi$. We may determine K_n^m by means of (29) and (30), though actual determination is made in the following sections.

Induction by the 24-hourly component of S_q . Since the inducing field of the 24-hourly component of S_q is given by (7), the current function at $r = a$ is to be expressed as

$$\begin{aligned} \Psi_1 = e^{i\alpha t} \left[\sum_n a_n P_n^1(\cos \theta) \cos \phi \right. \\ \left. + \sum_n b_n P_n^2(\cos \theta) \sin 2\phi \right] \end{aligned} \quad (31)$$

We see, therefore, by putting $p = i\alpha$, that conditions (29) and (30) lead to

$$\begin{aligned} - \sum_n n(n+1)a_n P_n^1 \\ = 2a^2 K i \alpha (e_{2,c}^{-1} - i e_{2,s}^{-1}) (A_2 + i B_2) P_2^1 \\ + 2i \xi^{-1} \sum_n \frac{n(n+1)}{2n+1} \\ \cdot (A_n + i B_n) a_n P_n^1 \text{ for } 0 < \theta < \pi/2 \\ - \sum_n n(n+1)a_n P_n^1 = 0 \\ \text{for } \pi/2 < \theta < \pi \end{aligned} \quad (32)$$

and

$$\begin{aligned} - \sum_n n(n+1)b_n P_n^2 \\ = 2a^2 K \alpha (e_{2,c}^{-1} - i e_{2,s}^{-1}) (A_2 + i B_2) P_2^2 \\ + 2i \xi^{-1} \sum_n \frac{n(n+1)}{2n+1} \\ \cdot (A_n + i B_n) b_n P_n^2 \text{ for } 0 < \theta < \pi/2 \\ - \sum_n n(n+1)b_n P_n^2 = 0 \\ \text{for } \pi/2 < \theta < \pi \end{aligned} \quad (33)$$

where

$$\xi = (2\pi a K \alpha)^{-1} \quad (34)$$

and

$$\begin{aligned} A_n + i B_n = 1 - q^{2n+1} \\ \cdot [1 - F_n(k^2 q^2 a^2) / F_{n-1}(k^2 q^2 a^2)] \end{aligned} \quad (35)$$

After multiplying by $P_n^1 \sin \theta$, we integrate (32) with respect to θ from $\theta = 0$ to $\theta = \pi$, where the first expression is used for $0 < \theta < \pi/2$ and the second one for $\pi/2 < \theta < \pi$. The equation then becomes

$$\left. \begin{aligned} & (N+1)a_N R_{NN} \\ & = a^2 K i \alpha (e_{2,c}^{-1} - i e_{2,s}^{-1}) (A_2 + i B_2) R_{2N} \\ & + i \xi^{-1} \sum_n \frac{n(n+1)}{2n+1} (A_n + i B_n) a_n R_{nN} \\ & - \xi N(N+1) \bar{a}_N R_{NN} \\ & + \sum_n \frac{n(n+1)}{2n+1} a_n^* R_{nN} = \frac{a e_{2,s}^{-1}}{2\pi} R_{2N} \\ & - \xi N(N+1) a_N^* R_{NN} \\ & - \sum_n \frac{n(n+1)}{2n+1} \bar{a}_n R_{nN} = \frac{a e_{2,c}^{-1}}{2\pi} R_{2N} \end{aligned} \right\} \quad (40)$$

which, by putting $a_n = \bar{a}_n + i a_n^*$, we obtain

$$\left. \begin{aligned} & (N+1) \bar{a}_N R_{NN} \\ & \sum_n \frac{n(n+1)}{2n+1} (B_n \bar{a}_n + A_n a_n^*) R_{nN} \\ & \frac{2}{\pi} (e_{2,s}^{-1} A_2 - e_{2,c}^{-1} B_2) R_{2N} \\ & (N+1) a_N^* R_{NN} \\ & \sum_n \frac{n(n+1)}{2n+1} (A_n \bar{a}_n - B_n a_n^*) R_{nN} \\ & \frac{2}{\pi} (e_{2,c}^{-1} A_2 + e_{2,s}^{-1} B_2) R_{2N} \end{aligned} \right\} \quad (36)$$

and

$$\left. \begin{aligned} & -\xi N(N+1) \bar{b}_N S_{NN} \\ & + \sum_n \frac{n(n+1)}{2n+1} \bar{b}_n^* S_{nN} = \frac{a e_{2,c}^{-1}}{2\pi} S_{2N} \\ & -\xi N(N+1) b_N^* S_{NN} \\ & - \sum_n \frac{n(n+1)}{2n+1} \bar{b}_n S_{nN} \\ & = -\frac{a e_{2,s}^{-1}}{2\pi} S_{2N} \end{aligned} \right\} \quad (41)$$

which are exactly the same equations as those obtained in the previous paper [Rikitake, 1960] for a single conducting sheet without the conducting mantle. R_{nN} and S_{nN} have been already given for $n \leq 6$ and $N \leq 6$ in that paper.

In a similar way, we multiply (33) by $P_N^2 \sin \theta$ and integrate it. The procedure leads to

$$\left. \begin{aligned} & (N+1) \bar{b}_N S_{NN} \\ & \sum_n \frac{n(n+1)}{2n+1} (B_n \bar{b}_n + A_n b_n^*) S_{nN} \\ & \frac{a}{2\pi} (e_{2,c}^{-1} A_2 + e_{2,s}^{-1} B_2) S_{2N} \\ & (N+1) b_N^* S_{NN} \\ & \sum_n \frac{n(n+1)}{2n+1} (A_n \bar{b}_n - B_n b_n^*) S_{nN} \\ & \frac{a}{2\pi} (-e_{2,s}^{-1} A_2 \\ & + e_{2,c}^{-1} B_2) S_{2N} \end{aligned} \right\} \quad (38)$$

Equations 36 are regarded as a set of simultaneous equations for the \bar{a}_n 's and a_n^* 's, and equations 38 give the \bar{b}_n 's and b_n^* 's. The factor ξ is estimated to be 0.859 for S_a provided the electrical conductivity of sea water is assumed to be 4×10^{-11} emu and the depth of the ocean is assumed to be 1000 m.² A_n and B_n can be estimated from (35). It is assumed that $\sigma = 5 \times 10^{-12}$ emu and $q = 0.94$ [Rikitake, 1950]. As for $e_{2,c}^{-1}$ and $e_{2,s}^{-1}$, the values given in Table 1 are used.

The simultaneous equations can be solved easily because the diagonal terms on the left-hand sides are fairly large. The solutions are shown in Table 2. In Table 3 the coefficients obtained by solving (40) and (41) are also shown. From Tables 2 and 3 it is apparent that the electric currents induced in the hemispherical sheet in the presence of the conducting mantle become several times smaller than those for the model without a conducting mantle.

When consideration is taken of the fact that the component of the electric currents normal

$$S_{nN} = \int_0^1 P_n^2(x) P_N^2(x) dx \quad (39)$$

we make $\sigma \rightarrow 0$, it is easily seen that $A_n \rightarrow 1$ and $B_n \rightarrow 0$. In this case (36) and (38) become

² In a future paper computations based on greater ocean depths will be given.

TABLE 3. The Coefficients for the Current Functions, in Gammas, Multiplied by $a/4\pi$ in the Hemispherical Ocean without the Conducting Mantle

n	\bar{a}_n	a_n^*	\bar{b}_n	b_n^*	$\bar{\alpha}_n$	α_n^*	$\bar{\beta}_n$	β_n^*	$\bar{\gamma}_n$	γ_n^*
1	-0.711	-4.929			-0.337	0.474				
2	-0.207	-2.587	-2.792	-0.214	-0.329	-1.349	-2.697	0.836		
3	0.045	-0.845	-1.360	-0.050	-0.223	-1.712	-2.001	0.487	-0.132	-1.3
4	0.050	-0.026	0.006	0.054	-0.148	-0.856	-0.787	0.102	-0.117	-0.7
5	-0.013	0.103	0.280	0.028	-0.034	0.013	-0.031	-0.027	-0.044	0.1
6	-0.021	0.010	-0.004	-0.024	-0.025	0.201	0.117	-0.002	0.019	0.1

to the ocean boundary should vanish at $\theta = \pi/2$, an alternative expression for the current function is obtained; that is,

$$\Psi_1 = e^{i\alpha t} \left[\sum_m c_{2m} P_{2m}^1(\cos \theta) \cos \phi + \sum_m d_{2m+1} P_{2m+1}^2(\cos \theta) \sin 2\phi \right] \tag{42}$$

which is true only for $0 < \theta < \pi/2$. On the other half of the spherical surface Ψ_1 is zero everywhere.

We therefore have

$$\left. \begin{aligned} \sum_n a_n P_n^1 &= \sum_m c_{2m} P_{2m}^1 \\ &\text{for } 0 < \theta < \pi/2 \\ \sum_n a_n P_n^1 &= 0 \text{ for } \pi/2 < \theta < \pi \end{aligned} \right\} \tag{43}$$

from which, by the use of multiplication and integration procedures similar to those adopted before, we obtain

$$2R_{NN} = \sum_m c_{2m} R_{2mN} \tag{44}$$

the solutions of which are

$$c_{2s} = 2a_{2s} \tag{45}$$

In a similar fashion we also obtain

$$d_{2s+1} = 2b_{2s+1} \tag{46}$$

The current systems induced in the ocean by the 24-hourly component of S_a can easily be obtained for epochs $\alpha t = 0^\circ$ and 90° by estimating values of c_{2s} and d_{2s+1} from the coefficients given in Tables 2 and 3. Such induced currents are shown in Figures 2 and 3 for the case in which the effect of the conducting mantle is taken into account. In these figures the current systems are viewed from a distant point above the center of the ocean. Currents of 2500 amp are flowing between adjacent lines, though it is

sometimes necessary to insert dotted lines which show the half values. These epochs correspond to the 0th and 6th hours at the $\phi_0 = 0$ meridian time. Current systems for epochs at the 12th and 18th hours can readily be obtained by reversing the direction of flow for those for the 0th and 6th hours.

The coefficients in Tables 2 and 3 are estimated on the assumption that the center of the ocean lies on the Greenwich meridian. If the center or $\phi_0 = 0$, meridian is assumed to be at $165^\circ W$, which may be taken as the approximate central meridian of the Pacific Ocean, it is easily seen that $\alpha t = 0^\circ$ and 90° correspond, respectively, to the 11th and 17th hours of Greenwich meridian time.

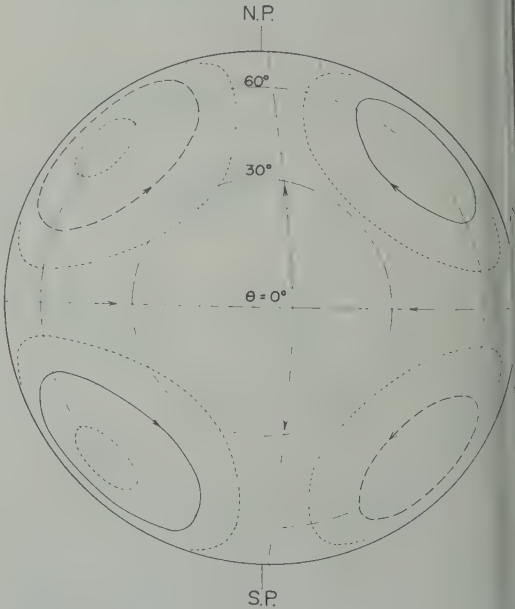


Fig. 2. Induced currents for $\alpha t = 0^\circ$ in the hemispherical ocean underlain by the conducting mantle by the 24-hourly component of S_a .

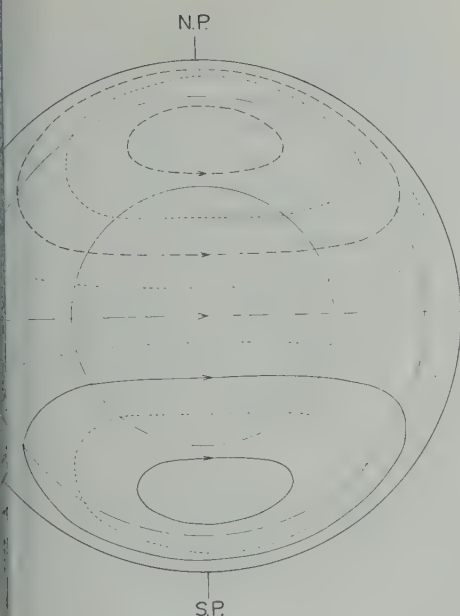


Fig. 3. $\alpha t = 90^\circ$.

duction by the 12-hourly component of S_a . Induction by the 12-hourly component of I_n can be studied in a way similar to that for the 24-hourly component. In this case the vector function is written as

$$e^{2i\alpha t} \left[\sum_n \alpha_n P_n^1(\cos \theta) \cos \phi + \sum_n \beta_n P_n^2(\cos \theta) \sin 2\phi + \sum_n \gamma_n P_n^3(\cos \theta) \cos 3\phi \right] \quad (47)$$

Simultaneous equations for α_n , β_n , and γ_n are

$$\left. \begin{aligned} & -\xi N(N+1) \bar{\alpha}_N R_{NN} \\ & + 2 \sum_n \frac{n(n+1)}{2n+1} (B_n \bar{\alpha}_n + A_n \alpha_n^*) R_{nN} \\ & - \frac{a}{2\pi} \frac{3\sqrt{10}}{4} (e_{3,s}^2 A_3 - e_{3,c}^2 B_3) R_{3N} \\ & - \xi N(N+1) \alpha_N^* R_{NN} \\ & + 2 \sum_n \frac{n(n+1)}{2n+1} (A_n \bar{\alpha}_n - B_n \alpha_n^*) R_{nN} \\ & + \frac{a}{2\pi} \frac{3\sqrt{10}}{4} (e_{3,c}^2 A_3 + e_{3,s}^2 B_3) R_{3N} \end{aligned} \right\} \quad (48)$$

$$\left. \begin{aligned} & -\xi N(N+1) \bar{\beta}_N S_{NN} \\ & + 2 \sum_n \frac{n(n+1)}{2n+1} (B_n \bar{\beta}_n + A_n \beta_n^*) S_{nN} \\ & = -\frac{a}{2\pi} 3(e_{3,c}^2 A_3 + e_{3,s}^2 B_3) S_{3N} \\ & - \xi N(N+1) \beta_N^* S_{NN} \\ & - 2 \sum_n \frac{n(n+1)}{2n+1} (A_n \bar{\beta}_n - B_n \beta_n^*) S_{nN} \\ & = \frac{a}{2\pi} 3(e_{3,s}^2 A_3 - e_{3,c}^2 B_3) S_{3N} \end{aligned} \right\} \quad (49)$$

$$\left. \begin{aligned} & -\xi N(N+1) \bar{\gamma}_N T_{NN} \\ & + 2 \sum_n \frac{n(n+1)}{2n+1} (B_n \bar{\gamma}_n + A_n \gamma_n^*) T_{nN} \\ & = \frac{-a}{2\pi} \frac{3\sqrt{6}}{4} (e_{3,s}^2 A_3 - e_{3,c}^2 B_3) T_{3N} \\ & - \xi N(N+1) \gamma_N^* T_{NN} \\ & - 2 \sum_n \frac{n(n+1)}{2n+1} (A_n \bar{\gamma}_n - B_n \gamma_n^*) T_{nN} \\ & = -\frac{a}{2\pi} \frac{3\sqrt{6}}{4} (e_{3,c}^2 A_3 + e_{3,s}^2 B_3) T_{3N} \end{aligned} \right\} \quad (50)$$

where

$$T_{nN} = \int_0^1 P_n^3(x) P_N^3(x) dx \quad (51)$$

and R_{nN} and S_{nN} are defined in (37) and (39). As before - and * show the real and imaginary parts of the coefficients, respectively.

The coefficients that are obtained as solutions of (48), (49), and (50) are also given in Table 2 for the present model and in Table 3 for the model without the conducting mantle. The coefficients so determined are used in drawing the current systems induced in the ocean by the 12-hourly component of S_a for epochs $\alpha t = 0^\circ$ and 45° . Figures 4 and 5 show the current systems for the present model. The electric current flowing between two adjacent lines is again 2500 amp. We again see that the induced currents are reduced by a factor 2 or 3 by the addition of the conducting mantle to the sheet.

Induced magnetic fields. When C_n is eliminated from (25) and (26), we obtain

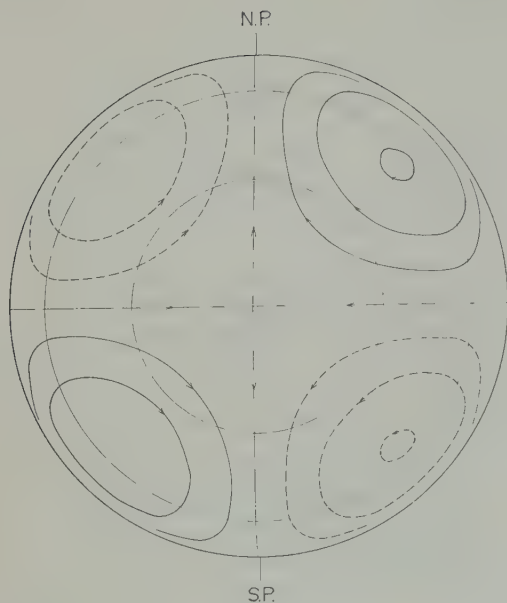


Fig. 4. Induced currents for $\alpha t = 0^\circ$ in the hemispherical ocean underlain by the conducting mantle by the 12-hourly component of S_q .

$$i_n^m = \frac{nq^{2n+1}}{n+1} [1 - F_n(k^2 q^2 a^2)/F_{n-1}(k^2 q^2 a^2)] e_n^m + \frac{4\pi}{a} K_n^m \frac{n}{2n+1} \{1 - q^{2n+1} \cdot [1 - F_n(k^2 q^2 a^2)/F_{n-1}(k^2 q^2 a^2)]\}$$

Since the first term on the right-hand side of the above expression gives the coefficient for the induced potential when there is the conducting part of the mantle only, the second term can be regarded as the effect of the hemispherical ocean on the induced potential. With the aid of (35), this effect is written as

$$\Delta i_n^m = (4\pi/a) K_n^m n / (2n+1) (A_n + iB_n) \quad (52)$$

The values of K_n^m are given in Table 2, so that Δi_n^m for each of the 24- and 12-hourly components of S_q is easily calculated. The real and imaginary parts of $\Delta i_n^m = \Delta i_n^m + \Delta i_n^{m*}$ are given in Table 4. This table enables us to illustrate the anomalous distributions of magnetic field and potential.

Because the coefficients given in Table 4 are so small that they amount to only several per cent of those of the inducing field, the anomaly caused by the presence of the ocean is also small.

The anomaly over the nonoceanic hemisphere is also extremely small, which is as it should be.

The distributions of the vertical magnetic field (positive upwards) for epochs $\alpha t = 0^\circ$ and 90° are calculated and shown in Figures 6 and 7 over only the oceanic hemisphere. It is clearly seen that the anomalous field is so small that its maximum amplitude does not exceed 2 gamma.

Discussion and conclusions. It turns out that the electric currents induced in a hemisphere of ocean by S_q are a few times smaller than those estimated for the model in which no account is taken of the electromagnetic coupling between the ocean and the conducting mantle. The maximum electric field in the sea associated with the induced currents is estimated to be a fraction of a mv/km.

The anomalous magnetic fields caused by the presence of the ocean, hardly exceeding 2 gamma at maximum amplitude, are unexpectedly small. The smallness of the anomaly is caused by two facts: (1) the induced currents in the ocean are smaller than hitherto supposed on the basis of a single sheet model and (2) the mantle covered by the ocean is shielded to some extent, resulting in a decrease in the induced currents there. The second fact, in turn, causes a marked decrease in the anomalous magnetic field at the surface of the ocean because of the reduction of

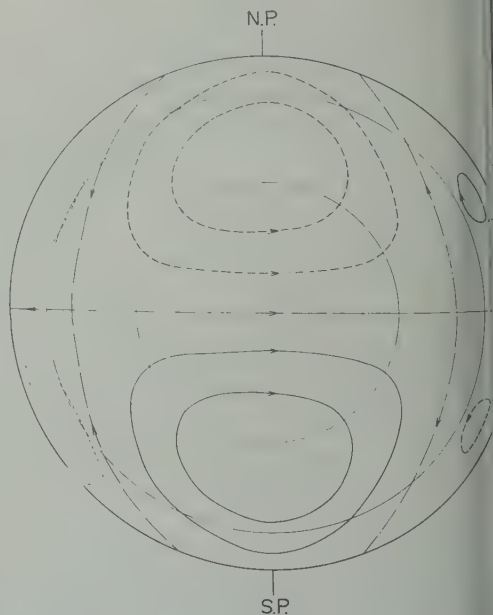


Fig. 5. $\alpha t = 45^\circ$.

E 4. Effect of the Ocean on the Magnetic al of S_q , in Gammas, 24-Hourly Component

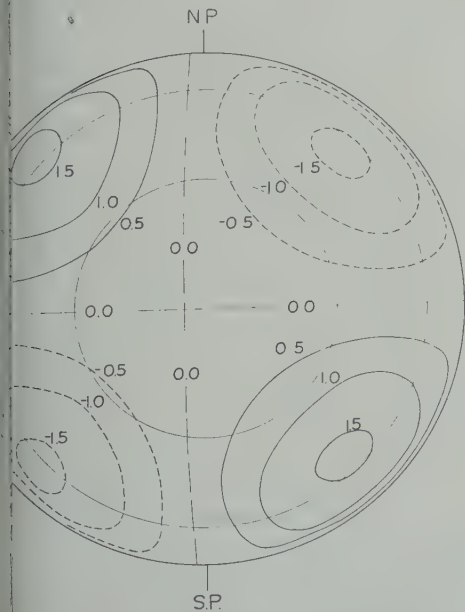
$\Delta \bar{z}_n^1$	$\Delta \bar{z}_n^2$	$\Delta \bar{z}_n^{1*}$	$\Delta \bar{z}_n^{2*}$
0.016		-0.137	
0.022	-0.132	-0.128	-0.006
0.004	-0.089	-0.056	0.006
0.000	0.001	0.000	0.004
0.000	0.027	0.010	0.002
0.000	0.000	0.000	0.000

12-Hourly Component

	$\Delta \bar{z}_n^2$	$\Delta \bar{z}_n^3$	$\Delta \bar{z}_n^{1*}$	$\Delta \bar{z}_n^{2*}$	$\Delta \bar{z}_n^{3*}$
05			0.004		
22	-0.187		-0.093	0.044	
21	-0.188	-0.020	-0.150	0.040	-0.118
10	-0.087	-0.015	-0.091	0.016	-0.081
02	0.000	-0.002	0.001	0.000	0.001
04	0.017	0.002	0.027	-0.002	0.026

roduced by the induced currents in the

pite of the long history of geomagnetic ation, nothing certain has been reported rd to the effect on S_q of a large ocean.



6. The ocean effect on vertical magnetic in units of gamma for $\alpha t = 0^\circ$ over the oceanic hemisphere.

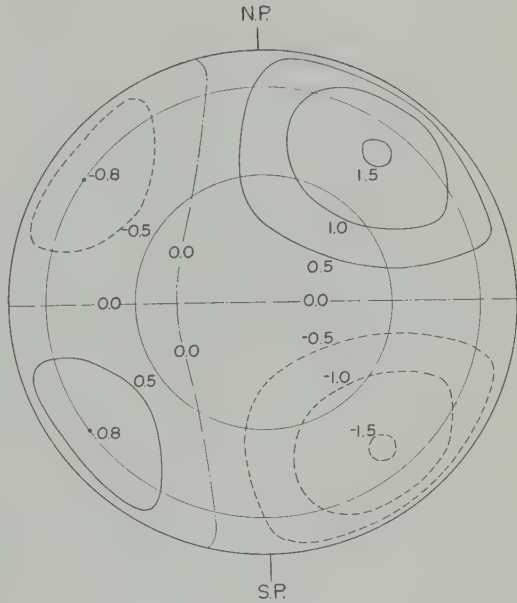


Fig. 7. The ocean effect on vertical magnetic field in units of gamma for $\alpha t = 90^\circ$ over the oceanic hemisphere.

Judging from what has been estimated in this paper, it seems not easy to detect the ocean effect on S_q because the anomalous magnetic fields amount to only a fraction of the inducing field. This seems to be one of the reasons why we have not been able to observe the effect clearly. Since the ocean effect depends on universal time, unlike the main part of S_q which is mostly dependent on local time, careful analysis with data from well-distributed observatories might be able to throw some light on the means of detecting the effect.

Acknowledgments. The writer thanks Prof. I. Özdoğan, the Director of the Geophysical Institute, Istanbul University, and Prof. R. Yazar, the Director of the Seismological Institute, Technical University of Istanbul, for their support.

REFERENCES

- Chapman, S., The solar and lunar variation of the earth's magnetism, *Phil. Trans., Roy. Soc. London*, 218, 1-118, 1919.
- Chapman, S., and J. Bartels, *Geomagnetism*, Oxford, 711-749, 1940.
- Chapman, S., and A. T. Price, The electric and magnetic state of the interior of the earth, as inferred from terrestrial magnetic variations, *Phil. Trans. Roy. Soc. London*, 229, 427-460, 1930.

- Lahiri, B. N., and A. T. Price, Electromagnetic induction in non-uniform conductors, and the determination of the conductivity of the earth from terrestrial magnetic variations, *Phil. Trans. Roy. Soc. London*, *237*, 509-540, 1939.
- Price, A. T., The induction of electric currents in non-uniform thin sheets and shells, *Quart. J. Mech. Appl. Math.*, *2*, 283-310, 1949.
- Rikitake, T., Electromagnetic induction within the earth and its relation to the electrical state of the earth's interior, *Bull. Earthquake Research Inst. Tokyo Univ.*, *28*, 45-100 and 219-283, 1950.
- Rikitake, T., Electromagnetic induction in a hemispherical ocean by S_4 , *J. Geomag. Geoelec.*, *11*, 65-79, 1960.
- Rikitake, T., The effect of the ocean on rapid magnetic changes, *Geophys. J.*, *3*, in press.
- Rikitake, T., and I. Yokoyama, The anomalous behaviour of geomagnetic variations of short period in Japan and its relation to the crustal structure, the 6th paper, *Bull. Earthquake Research Inst., Tokyo Univ.*, *33*, 297-331, 1957.
- Satô, Y., Transformation of wave-function related to the transformation of coordinate systems, *2*, *Bull. Earthquake Research Inst. Tokyo Univ.*, *28*, 175-217, 1950.

(Manuscript received May 8, 1961;
revised July 26, 1961.)

Backscattering of 3.21-Centimeter Radiation by Water Bubbles

LOUIS J. BATTAN AND BENJAMIN M. HERMAN

*Institute of Atmospheric Physics, The University of Arizona
Tucson, Arizona*

Abstract. Calculations have been made of the backscattering of 3.21-cm waves from water bubbles ranging in diameter, D , from 0.2 to 5.0 cm. Film thicknesses between 2 and 10^3 microns were considered. As the diameter increases, the backscattering cross section shows a gradual upward trend, but sharp minima occur at intervals of D that are at integral values of $\lambda/2$. As the film thickness increases there is a general increase of the backscattering cross sections, the increase being most pronounced at small film thicknesses. Except at intervals of D close to $\lambda/2$, the backscattering cross sections from bubbles are higher than those from water spheres of the same mass.

Introduction. A great deal of attention has been given in the last few years to the backscattering of microwaves by water and ice spheres. The authors have recently completed a series of calculations of the backscattering cross sections of dry and wet ice spheres [Herman and Battan, 1959, 1960]. Equations developed by Kerker and Kerker [1951, 1952] from the complete Mie equations were used for this purpose. Aden's [1951] theory deals with concentric spheres having different indices of refraction. In studying the backscattering from a water bubble surrounding a sphere of ice. In studying the backscattering cross sections of a bubble, we have a 'sphere of air' surrounded by a shell of

ice. This problem is of interest for several reasons. The results of these calculations assist in the interpretation of the results of calculations and measurements of the backscattering by wet ice spheres. Atlas, Harper, Ludlam, and Macklin [1959] and Gerhardt, Tolbert, and Brunstein [1960] have recently completed a series of measurements which, in general, are supported by these calculations.

Fournier d'Albe in private communication proposed the possibility of seeding clouds with ice crystals. (Some aspects of the proposed new seeding technique are mentioned in a recent technical article by Fournier d'Albe [1960].) It was suggested that radar might be used to detect the bubbles and note whether they lead to rain. In this sense the bubbles may act as nuclei as well as seeding agents.

Results. An electronic computer (IBM-650) was used to calculate the backscattering cross

sections of bubbles whose diameters ranged from 0.2 to 5.0 cm and whose film thickness ranged from 1 to 8 microns. In a few cases greater film thicknesses were considered. The bubbles were taken to be composed of water with a complex index of refraction of $7.14-2.44i$, a value appropriate at a wavelength of 3.21 cm and a temperature of 0°C .

Figure 1 shows a series of curves of backscattering cross section σ as a function of film thickness F for various bubble diameters. It is seen that σ increases monotonically with F . Until the film thickness reaches about 0.005 cm, σ increases at about the same rate for the three drop sizes shown. At larger values of F the slopes of the curves differ considerably as the particles behave more and more like homogeneous water spheres.

Figure 2 is a plot of σ as a function of bubble diameter D for various film thicknesses. The extremely large changes of σ over small intervals of D are striking. It seems clear that these are interference patterns. The three peaks of σ are separated by intervals of D of about 1.5 and 1.6 cm; the three minima are separated by diameter differences of 1.7 cm. The average of these four intervals is 1.6 cm, a value very close to one-half a wavelength. The differences of the individual readings have come about, at least in part, because calculations were made of intervals of D of 0.2 cm.

It is clear that the backscattered wave is composed of energy scattered from all four interfaces, two air-water and two water-air. The resulting backscattered beam is then the superposition of many multiscattered waves of various

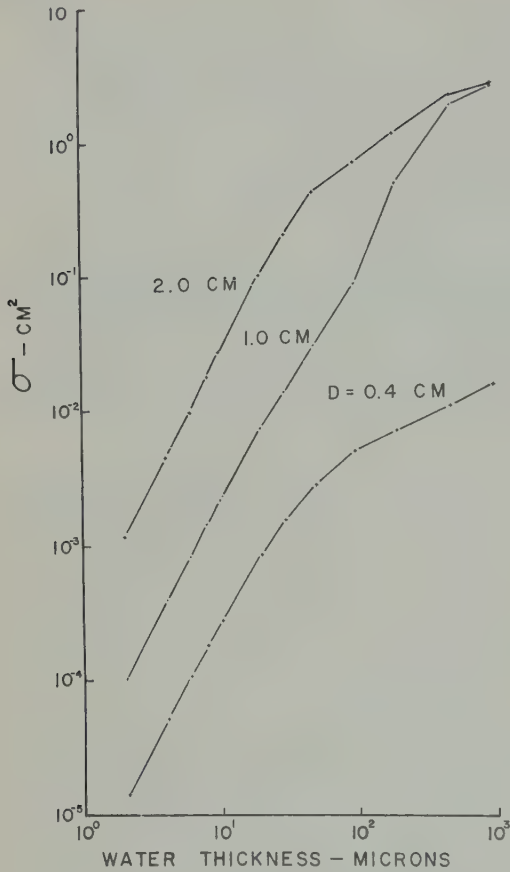


Fig. 1. Backscattering cross sections as functions of film thickness for bubbles having diameters of 0.4, 1.0, and 2.0 cm ($\lambda = 3.21$ cm). Water spheres with these diameters have backscattering cross sections of approximately 0.03, 2.03, and 2.04 cm², respectively.

phase relationships, and so it becomes difficult to interpret the resulting interference patterns by the techniques of geometrical optics. But because the maxima and minima occur at intervals of diameter $D = \lambda/2$ it appears highly probable that they are due to interference between the rays scattered from the two front interfaces and those scattered from the two rear interfaces.

From an examination of Figures 1 and 2 three important points can be made: with film thicknesses of the order of a few microns the backscattering cross sections are two to three orders of magnitude smaller than with all-water drops and the backscattering increases with film

thickness; the variations of σ with D are such as to produce minima at intervals of $\lambda/2$; and the phase of the amplitude is such that minima occur at values of D corresponding to $n\lambda/2$ and maxima at $D = [(2n - 1)\lambda]/4$, where n is an integer.

We have not succeeded in arriving at a satisfactory physical explanation of all three observations.

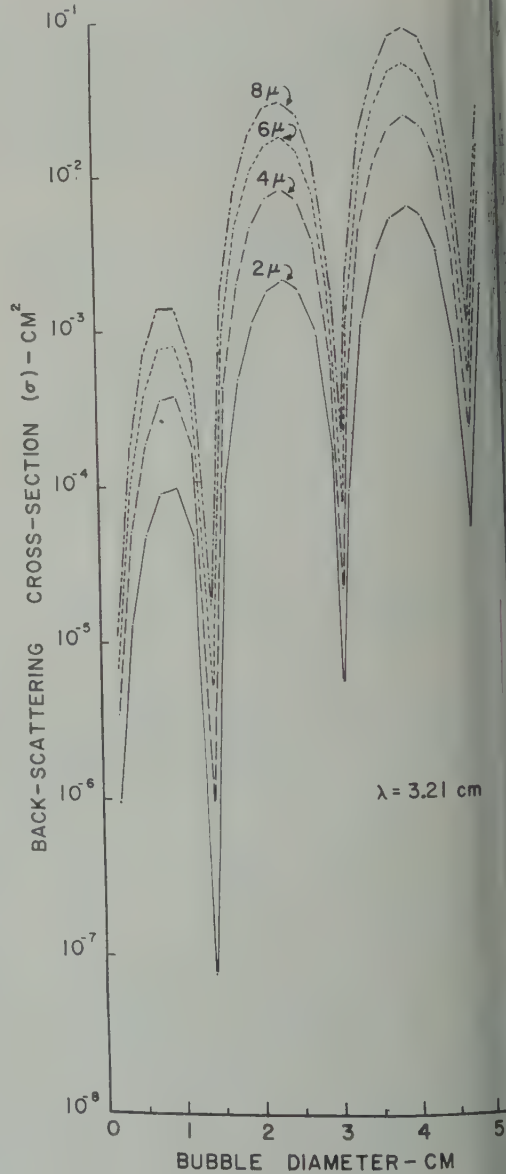
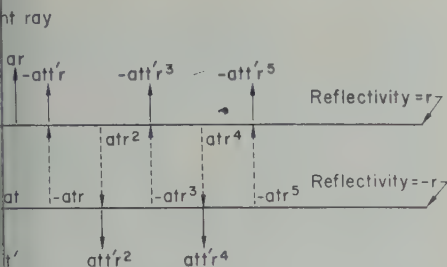


Fig. 2. Backscattering cross sections of bubbles as functions of bubble diameter for film thicknesses indicated on the curves.



3. Multiple reflections of a ray in a thin layer of water.

but we believe that the first can be explained in terms of the interference between rays from the front and back surfaces of the bubble film. Apparently the backscattering from the bubbles has a small net amplitude because of many scattered waves, most of which almost cancel each other. Let us consider waves scattered from the front two interfaces under the assumption that it was a flat surface. (See Fig. 3.) Here, the reflectivity of the water-air interface is r , that of the water-air interface is $-r$ because of the phase change, the transmission from air to water is t , and that from water to air is t' , where $tt' = 1 - r^2$. Now, if the thickness of the water film is $\lambda/2$, the total distance traveled by the internally reflected rays is some integral multiple of λ . Then it can be shown by summing all the reflected rays except the first that they add up to $-ar$. They exactly cancel the first reflected ray, and the resulting intensity is zero. When the film thickness equals $\lambda/4$ the backscattering has a maximum value. As the film thickness decreases from $\lambda/4$ to 0, the backscattering increases and approaches zero. With a film thickness of several microns each reflected ray is out slightly out of phase relative to the incident ray. Cancellation is not complete, and the result is a small net backscattered ray. In a real bubble, of course, to the net amplitude of the scattered beam from the front film must be added the waves resulting from multiple reflections within the back film, plus reflections from beams scattered back and forth between the two films. It is this last group of reflections that cause the diameter dependence of the backscatter, and their total distance traveled, and therefore their phase, will depend on the diameter of the bubble. Thus, the net reflection is the superposition of many waves which nearly, but

not quite, cancel each other. As the film thickness increases, absorption of the reflected rays increases, thereby canceling less of the primary reflected beam, resulting in an increase of σ toward the all-water value and a decrease of the amplitude of the interference effects.

Although the principles of geometrical optics appear to explain the reduction of σ with increasing F , the location of the minima cannot be explained in this way. If the bubble is considered to consist of two parallel films separated by a distance D , we would expect to find a minimum at $\lambda/4$, but obviously this does not occur. This aspect of this problem needs further attention.

Atlas, Harper, Ludlum, and Macklin [1960] have suggested that an all-water sphere acts like a metal sphere in the sense that all the backscattering is from the front surface. They have stated that this is true because virtually all the incident electromagnetic energy is either scattered or absorbed in the first 100 microns (approximately) of water. The calculations in Figure 1 show that this conclusion is not quite correct for some of the larger particles. For example, the increase of σ from a sphere of 1.0-cm diameter is very large when the film thickness increases from 0.01 to 0.05 cm. It is clear, however, that most of the increase of σ does occur before F reaches 0.01 cm.

Radar detectability of bubbles. If bubbles are to be used as radar tracers, we must know whether their reflectivities are high enough to give radar echoes. For the sake of this comparison, bubbles are assumed to have a film thickness of 4 microns. Ranz [1959] concluded, from experimental work, that a film thickness of 4 microns is reasonable.

One point of interest is how the σ of bubbles compare with water sphere of the same mass.

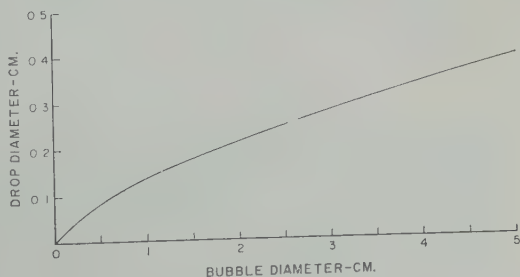


Fig. 4. Diameter of drops having the same mass as bubbles of the indicated mass.

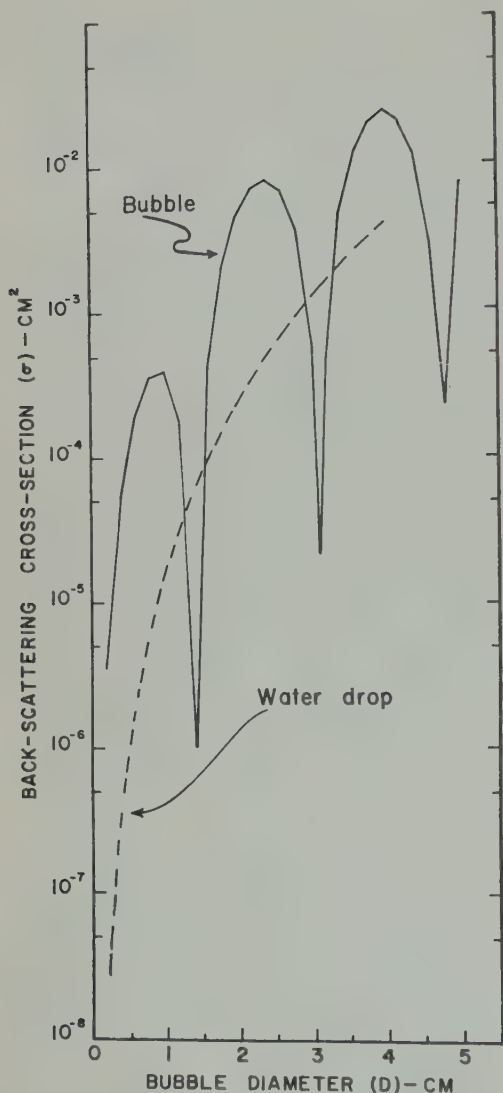


Fig. 5. Backscattering cross sections as functions of bubble diameter. Solid curve, bubble of 4-micron film thickness; dashed curve, water drop having same mass as a bubble of the diameter indicated on the abscissa.

Diameters of these equivalent spheres as a function of bubble diameters are shown in Figure 4. The backscattering cross sections of the homogeneous spheres were calculated from the Rayleigh approximation

$$\sigma = \frac{\pi^5}{\lambda^4} \left| \frac{m^2 - 1}{m^2 + 2} \right|^2 D^6$$

TABLE 1. Calculated Values of $\Sigma\sigma$ Needed for Radar Echo at Various Ranges and the Number of Bubbles of 1-Centimeter Diameter Needed for Volumes of the Given Size

Range, km	Volume, m ³	$\Sigma\sigma$, cm ²	Number of Bubbles
5	1.3×10^6	7.0×10^{-1}	1.8×10^1
10	5.0×10^6	1.1×10	2.9×10^1
20	2.0×10^7	1.8×10^2	4.7×10^1
30	4.5×10^7	9.1×10^2	2.4×10^2

where λ is wavelength, m is the complex index of refraction, and D is diameter. Over the range of particle sizes involved, this approximation is quite adequate.

It is evident from Figure 5 that, over many sizes considered, the bubbles have higher reflectivities than water drops of the same mass. The reverse is true when the bubble diameters are close to $(n/2)/\lambda$, where n is any integer. On the other hand, the bubble reflectivities are high at diameters close to $[(2n - 1)/4]\lambda$, where n is any integer. To detect water bubbles with radar, bubble diameters close to these values should be chosen.

The backscattered power from a volume of randomly distributed bubbles is given by

$$\bar{P}_r = \frac{P_t A_p^2 \sum \sigma}{9\pi\lambda^2 r^4}$$

where P_t and P_r are the peak transmitted and received power respectively, A_p is the aperture area of the antenna, and r is the range. The summation is performed over all the bubbles from which power is scattered back to the radar set so as to arrive at the same instant. Also the averaging of P_r must be performed over a sufficient period of time (of the order of 10 sec) to allow the bubbles to array themselves into a number of independent distributions.

Taking AN/TPS-10A as a typical radar set, we may calculate the number of bubbles needed to give an echo. The properties of this radar set are: $P_t = 65$ kw; $\lambda = 3.3$ cm; $A_p = 2.05$ m²; and the minimum detectable signal is about 10⁻⁸ watt.

Table 1 shows the results of the calculation of the number of bubbles of 1-cm diameter required to just give an echo at the indicated ranges. Since the bubbles must backscatter rad

that arrives at the radar at the same time, the number of bubbles calculated must be within a volume defined by a range interval to half the radar pulse length and the widths of the radar set. These volumes, at various ranges, are shown in Table 1. For the FPS-10, half the pulse length amounts to 1500 ft and the beam widths are 2.0° and 0.7° in horizontal and vertical planes respectively. Using the values in the table and Figure 2, and the requirements set forth above, the numbers of bubbles needed to give an echo can be calculated.

In summary the authors wish to point out that this study was not initiated for the purpose of promoting a program for cloud seeding with bubbles. Rather, the aim has been to learn something about the backscattering properties of water bubbles and to determine whether they have greater radar reflectivities than water drops of the same mass. For most bubble sizes, the answer is yes. If bubbles of approximately $(1/4)\lambda$ could be produced, the radar reflectivities would be more than an order of magnitude greater than water drops of the same

size. Finally, it should be noted that this study considered only bubbles composed of water. If films of soap or glycerine were considered the results would be slightly different because the index of refraction of the liquid would be different, although the essential features of the results would not be expected to be different.

Acknowledgments. We wish to thank General S. R. Browning (U.S.A. ret.) of the Numerical Analysis Laboratory of The University of Arizona for his contribution to this work.

This research has been supported by the Office of Naval Research under contract NR-032-164 and the National Science Foundation under grant G-8216.

REFERENCES

- Aden, A. L., and M. Kerker, Scattering of electromagnetic waves by two concentric spheres, *J. Appl. Phys.*, **22**, 1242-1246, 1951.
- Aden, A. L., Back-scattering of electromagnetic waves from spheres and spherical shells, *Geophys. Research Paper 15*, Geophys. Research Directorate, A. F. Cambridge Research Center, Bedford, Mass., 1952.
- Atlas, D., W. G. Harper, F. H. Ludlam, and W. C. Macklin, Radar scatter by large hail, *Dept. Meteorol. Imp. Col. Sci. Technol. Tech. (Sci.) Note 2*, London, 1960.
- Fournier d'Albe, E. M., Why not seed clouds with beer? *Weather*, **15**, 243-245, 1960.
- Gerhardt, J. R., C. W. Tolbert, and S. A. Brunstein, and W. W. Bahn, Experimental determinations of the back-scattering cross-sections of water drops and of wet and dry ice spheres at 3.2 centimeters, *J. Meteorol.*, **18**, 340-347, 1961.
- Herman, B. M., and L. J. Battan, Calculations of Mie back-scattering of microwaves from ice spheres, *Quart. J. Roy. Meteorol. Soc.*, **87**, 223-230, 1961.
- Herman, B. M., and Louis J. Battan, Calculations of Mie back-scattering from melting ice spheres, *Inst. Atmospheric Phys. Univ. Ariz. Sci. Rept. 15*, Tucson, 18 pp., 1960.
- Ranz, W. E., Some experiments of the dynamics of liquid films, *J. Appl. Phys.*, **30**, 1950-1955, 1959.

(Manuscript received May 22, 1961.)

On the Vertical Circulation of the Mediterranean Sea¹

GEORG WÜST

*Institut für Meereskunde at the University of Kiel
and*

*Lamont Geological Observatory, Columbia University
Palisades, New York*

Abstract. In this analysis of about 600 stations (200 winter and 400 summer stations) the first attempt is made to study, with the help of the 'core method,' the mean steady state of the deep circulation within the whole expanse of the Mediterranean, divided by sills into eight basins. In spite of the aperiodic fluctuations, we have got some indications of the seasonal variations of the Levantine intermediate current, which has its maximum in the winter. At the surface, six water types are formed which spread out, either by measurable currents or by weak advective processes, in six core layers and cause renewal and ventilation all the way to the bottom of the basins. On the whole, the Mediterranean vertical circulation offers, by the transformation of the entering Atlantic water type to the Mediterranean types, an excellent example of interaction between atmosphere and sea. This is demonstrated by a three-dimensional block diagram of the vertical circulation and of the salinity distribution during winter.

By its intercontinental situation in the midst of a subtropical semiarid climate and by its morphological structure, the Mediterranean Sea is a unique field for the application of the so-called 'core method.' This method permits to follow the spreading and mixing processes of the main water masses along their curved core layers, characterized by intermediate maxima and minima of salinity, oxygen, and temperature. In this way the main features of the mean steady state circulation can be delineated in the whole expanse of this enclosed sea and for different seasons.

By means of the vertical distribution of salinity, oxygen, and temperature we can identify four different core masses: (1) The near-surface water of Atlantic origin between 0 and 75 m depth, (2) the intermediate water between 200 and 600 m, (3) the deep water between 1500 and 3000 m, and (4) the bottom water at depths to 4200 m. I should like to demonstrate with four maps and four diagrams the main results of our recent study on the deep circulation for which about 600 stations (about 200 winter and 400 summer stations) of 12

research vessels are available [Wüst, 1960, 1961].

The first core map (Fig. 1) concerns the Levantine intermediate water, characterized by the maximum of salinity, which, except for the source region in the northern Levantine basin, is found in the whole Mediterranean at various morphologically influenced depths between 200 and 600 m. It is formed in February and March on both sides of Rhodes, where at the surface there is a combination of low temperatures (about 15°) and high salinities (39.1‰), i.e., conditions favorable for a vertical thermal-haline convection reaching to a depth of about 100–200 m. From this winter source region of high salinity the Levantine intermediate water spreads out within the core layer to all western basins.

After having passed the central Ionian basin the main flow goes over the Sicilian ridge through the Strait of Sardinia and along the continental slope of North Africa. This flow attains more and more the character of a real boundary current with measurable velocities, which we call the Levantine intermediate current. It finds its continuation in the outgoing undercurrent through the Strait of Gibraltar where this undercurrent in 275-m depth reaches the high velocities of more than 100 cm/sec.



Fig. 1. Distribution of salinity within the core layer of Levantine intermediate water in winter.

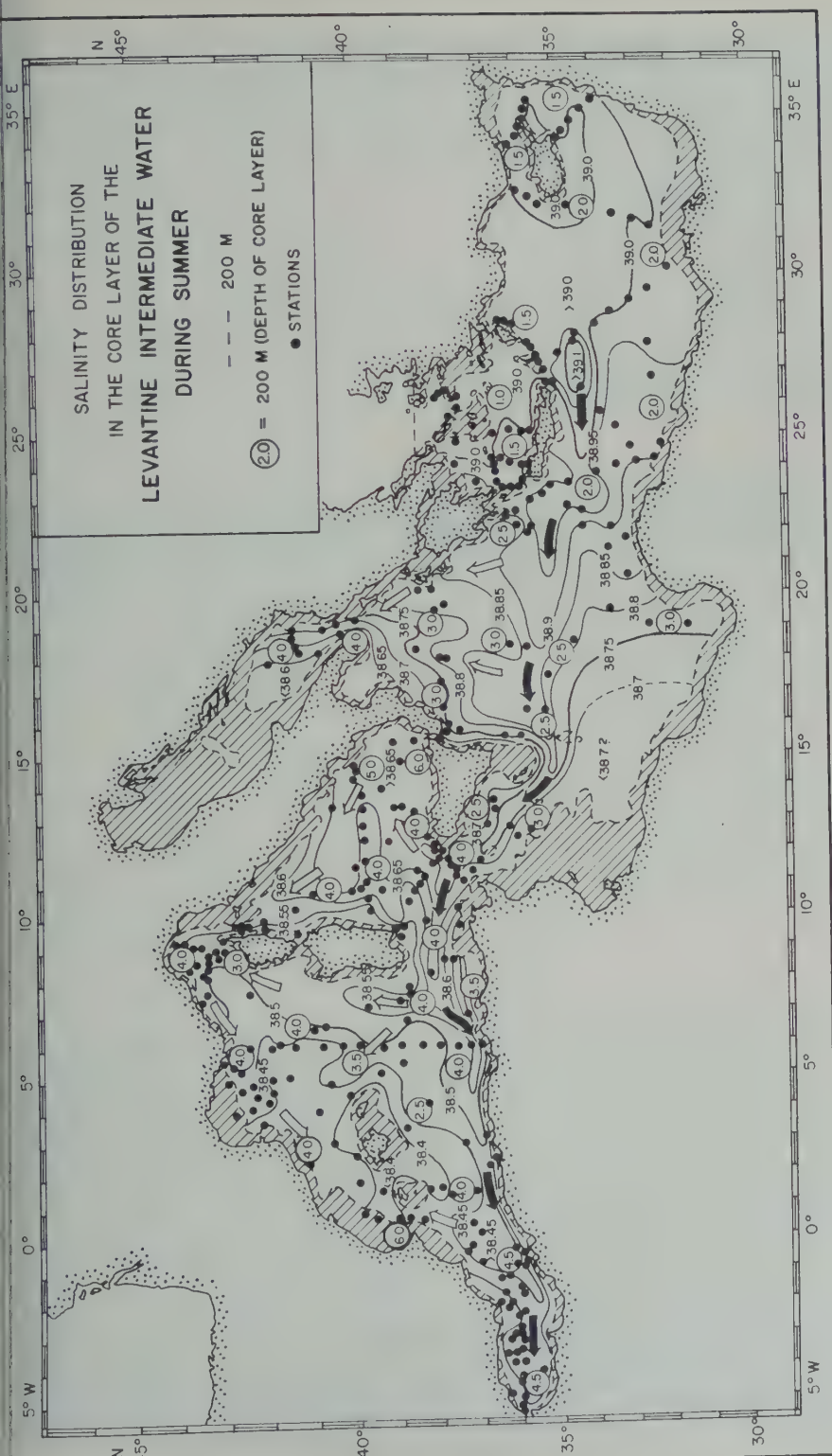


Fig. 2. Distribution of salinity within the core layer of Levantine intermediate water in summer.

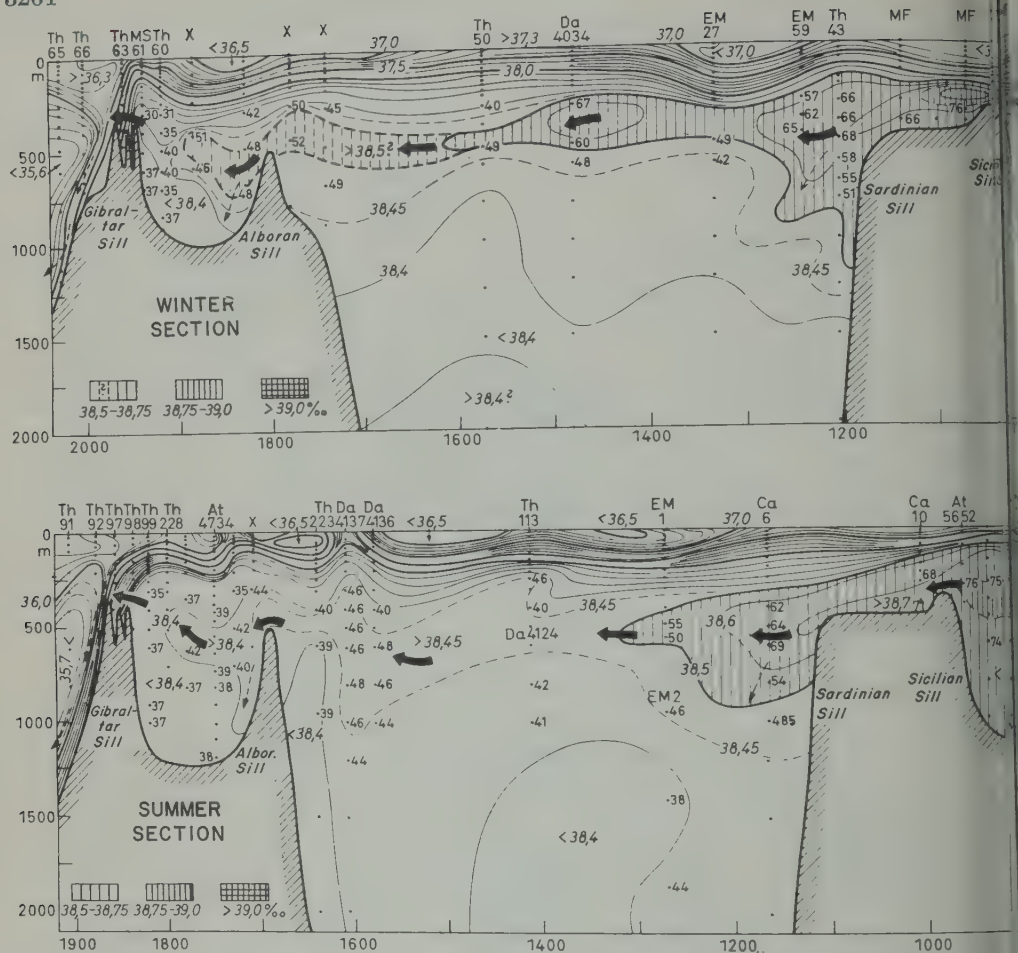


Fig. 3. Longitudinal sections of salinity along the axis of flow.

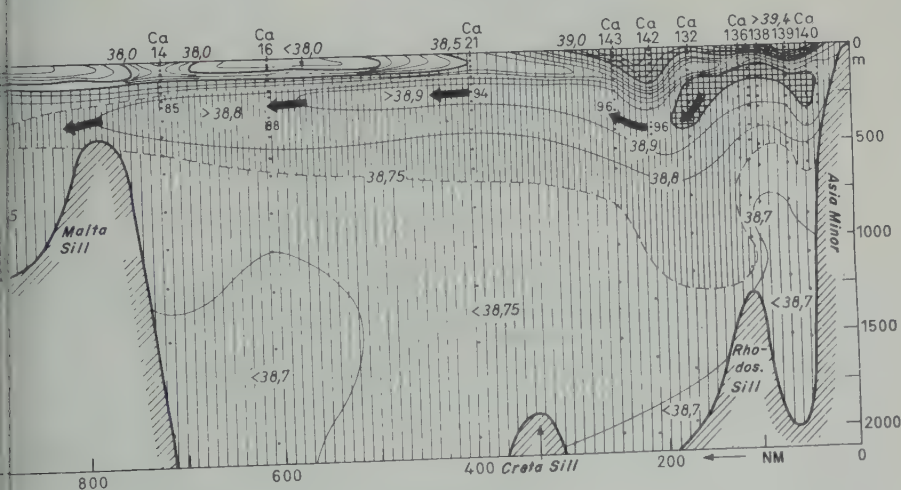
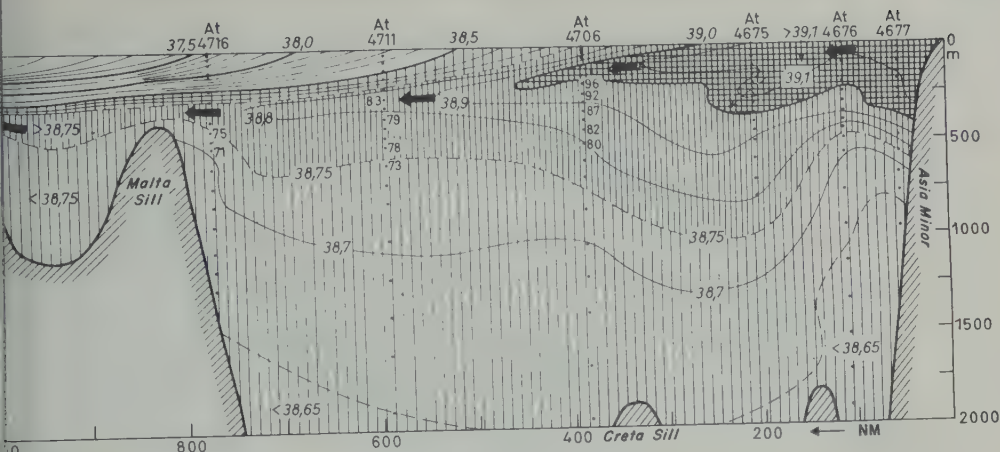
In the summer the Levantine intermediate current is perceptibly weaker than in winter (Fig. 2). But apart from this fact the main trends of the distribution of the salinity within the core layer remain the same in summer as in winter, which confirms the belief that the inhomogeneities in the observation material and the aperiodic fluctuations of salinity and other disturbing effects are of second order of magnitude.

The fact of a seasonal variation in the strength of the Levantine intermediate current is clearly demonstrated by the two longitudinal sections of salinity along the main axis of its flow, as shown in Figure 3. The near-surface Atlantic

undercurrent to the east² and the Levantine intermediate current to the west represent the most important branches of the Mediterranean deep circulation. The spreading and the mixing processes of the latter can also be described by means of *T/S* curves for winter and summer, as shown in Figure 4.

One end point of the winter normal curve represents the conditions at the source region

² The Atlantic undercurrent, characterized by intermediate salinity minimum, is well developed only during the summer at 20- to 75-m depth, chiefly along the North African continental shelf, as is clearly shown by the core map of Lacombe and Tschernia [1960, Fig. 3].



Levantine intermediate current for winter and summer.

Rhodes (point A). The other end points A and B^2 characterize the temperature, salinity, and density in the region, where the last traces of the Levantine water type disappear, that is, at the westernmost borders of the Balearic basin. It is remarkable that the T/S curve $A-B_2$ practically coincides with the σ_t line of 29.06, which indicates 'isentropic' advection or current along a density surface. In the winter there is a deviation of the density surface, but it is doubtful whether such a small difference can be related to a stronger circulation and mixing effects.

Another outstanding circulation problem concerns the formation and spreading of the

deep and bottom waters in the western and in the eastern Mediterranean, separated by the Sicilian sill (330 m). The enhancement of back radiation and loss of sensible heat accompanying invasions of polar air masses from the northern continent during February and March produce relatively cold (10° to 12.5°C) and heavy ($\sigma_t = 29.1$ to 29.2) water masses, formed at the surface of the following northern border regions: (1) Golfe du Lion and the areas along the Catalonian and French Riviera, (2) Ligurian Sea, (3) the southern Adriatic Sea, and (4) the southern Aegean Sea (probably of minor importance).

In the whole water column (0 to 1000 m) of the northern Balearic basin, there exists in

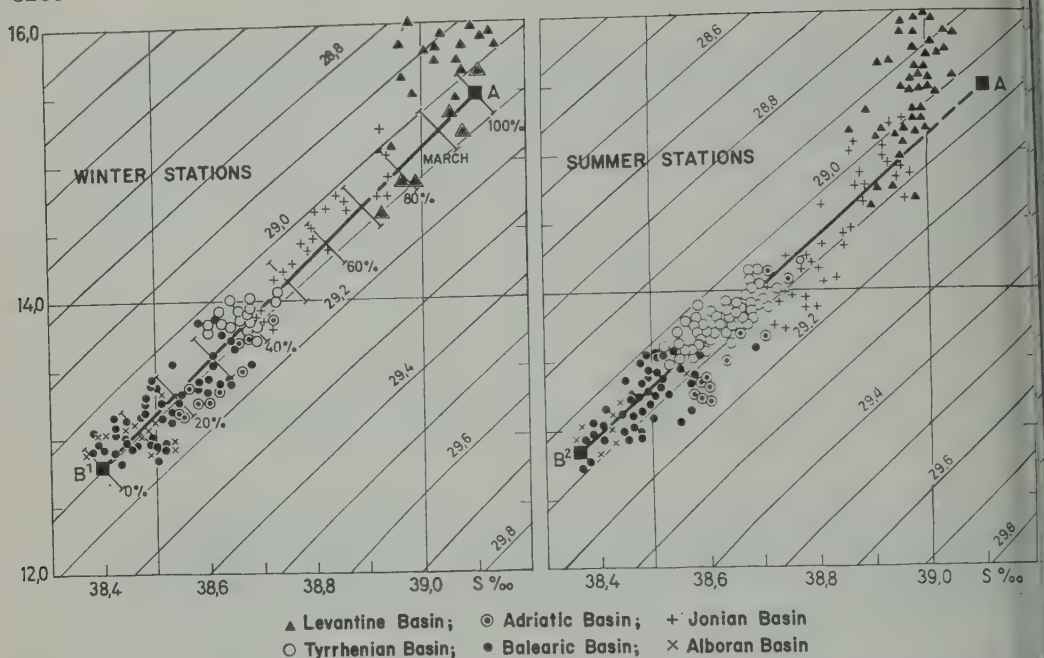


Fig. 4. T/S diagram for the core layer of the Levantine intermediate water in winter and summer.

February and March a nearly homogeneous vertical density distribution which favors extensive vertical convection. The mixing of surface and intermediate waters forms deep and bottom water of relatively high oxygen and high salinity which we call the *North Balearic deep water*.

This water type, by its spreading at depths between 1500 and 3000 m in the Balearic basin, produces an oxygen maximum of 4.4 to 4.7 cm^3/liter . In the Tyrrhenian basin, which is separated from the Balearic by a sill 1500 m deep, the ventilation of the deep water results in only 4.3 cm^3/liter (Fig. 5).

In the eastern Mediterranean the main source region of high oxygen deep water is situated in the southern Adriatic Sea, where by vertical convection and mixing of surface and saline intermediate water a bottom water of very high oxygen ($>5.0 \text{ cm}^3/\text{l}$) and density ($\sigma_t = 29.2$) is formed in February and March. This *Adriatic deep and bottom water* has, as shown by Pollak [1951], 'sufficiently high density to sink to those depths of the eastern Mediterranean at which the deep oxygen maximum is found.' We agree also with Pollak that 'there is little reason to doubt that the deep outflow through the Strait of Otranto is far from continuous.' But with

regard to the deep circulation in the Ionian and Levantine basins we cannot find 'a counter-clockwise circulation,' as Pollak has assumed, but deduce a main eastward spreading through the deeper regions of the two basins.

The *bottom waters* are best characterized by the potential temperature, salinity, and potential density of the nearest bottom layer. But we have to take into account the fact that most of the near-bottom observations are accomplished during serial measurements which normally end at various distances (100 m or more) from the bottom. In spite of this fact we have in the T_p/S diagram (Fig. 6), because of the vertical homogeneity, a modest spread of the points with standard deviations of only ± 0.03 per mil salinity.

Because of the separation by the Sicilian ridge, we find in the eastern Mediterranean different characteristics from those in the western, where (apart from the Adriatic Sea) densities, salinities, and temperatures are lower. In both cases, however, the normal curves are practically linear and parallel, which is to say that they follow the same linear mixing rule. It is remarkable that they cross the density lines. Therefore, contrary to the T/S relation

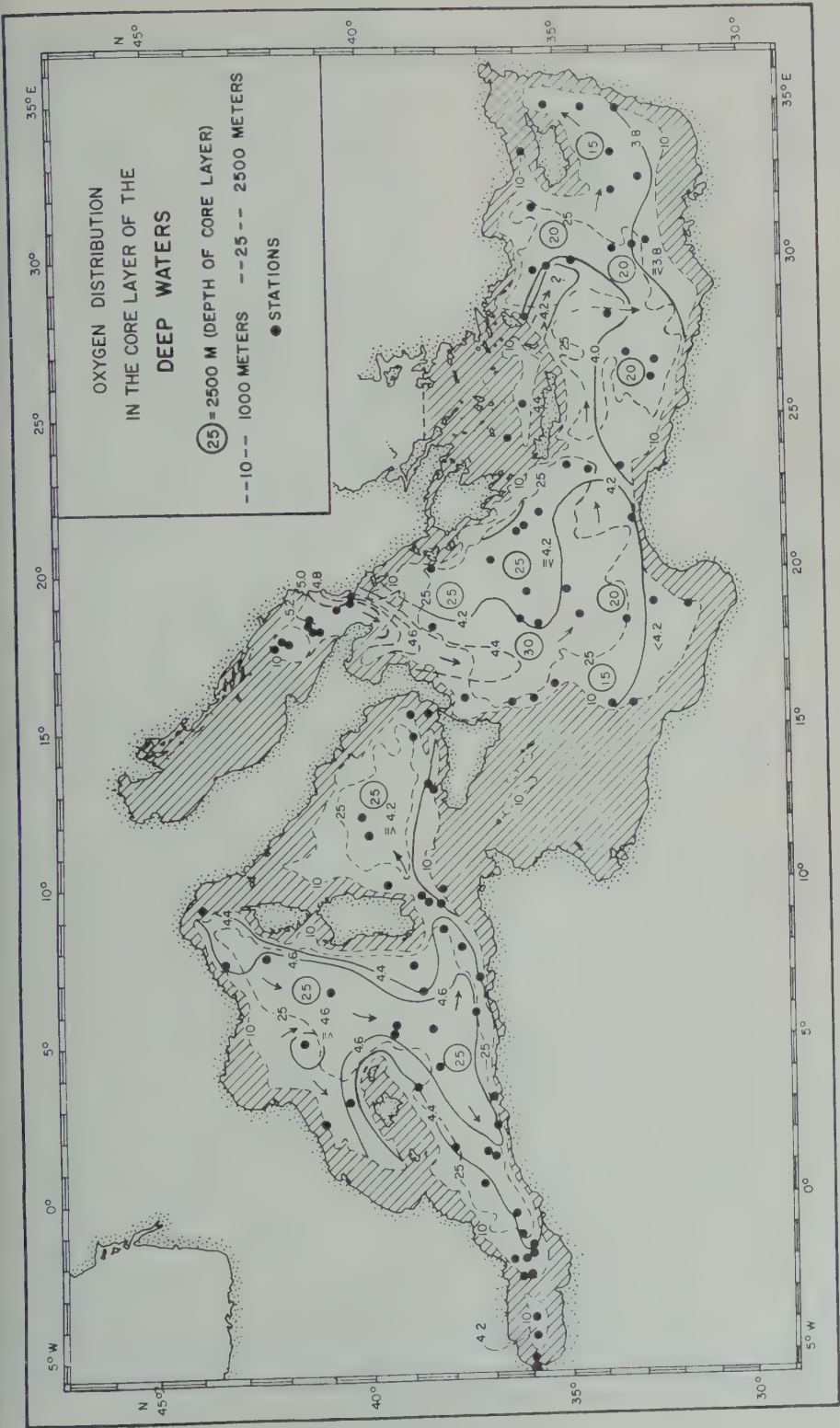


Fig. 5. Distribution of oxygen within the core layers of the deep waters.

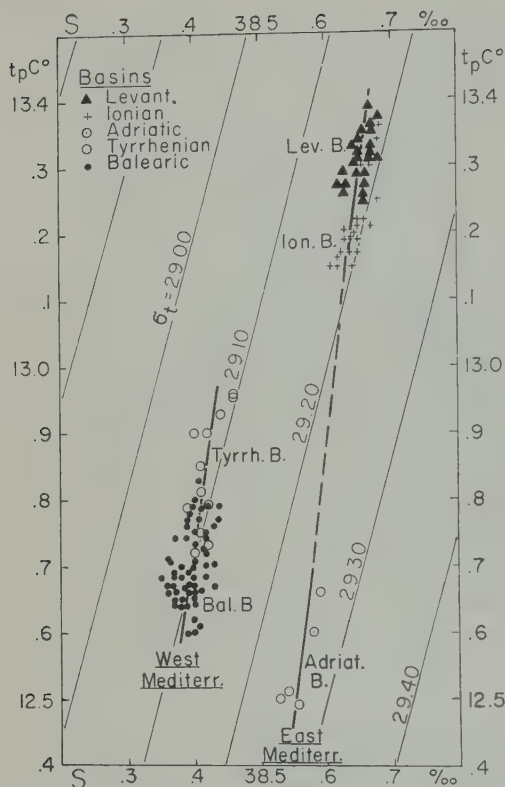


Fig. 6. T_p - S diagram within the bottom water (> 1500 m).

ship within the Levantine intermediate water (Fig. 4), in the North Balearic and the Adriatic bottom waters we do not have an isentropic advection along a special density surface, but stronger mixing effects and an oblique weak advection along the slopes of the bottom.

In the map of potential temperature of bottom water (Fig. 7) the interpretation with regard to the main flow of the various water masses is marked by small arrows. In the deepest areas of the Balearic and Tyrrhenian basins we find a very uniform distribution of temperature (12.65° and 12.74°C, respectively); at the borders there is an oblique temperature gradient, which means that the isotherms are here nearly parallel to the contour lines of depth.

The main flow of the *Adriatic bottom water* (of originally 12.6°C) is at first directed to the south and accompanied by marked vertical mixing on the Otranto sill. Therefore it attains temperatures of 13.15°C at the greatest depths

(>4000 m) in the central regions of the Ionian basin. From here, similar to the deep water, it turns to the east and after crossing the sill depth in the middle of the Crete sill it enters the Levantine basin where it flows with nearly constant temperatures of 13.27°C into the great depths south of the Rhodes sill and of Cyprus.

From the two maps of deep and bottom water it seems probable that some smaller influences come from the Aegean Sea by occasional overflow through the channels between Crete and Rhodes. But because of the small number of observations, the conditions of this overflow cannot yet be sufficiently examined.

Our analysis is a first attempt to study with the help of the core method the mean steady state of the deep circulation within the whole expanse of this enclosed sea, divided into eight deep basins. In spite of the aperiodic fluctuations and other disturbing influences, which are apparently of second order of magnitude, we have got some indications of the seasonal variations of the Levantine intermediate current which has its maximum in the winter. On the whole the vertical circulation offers by this transformation of the Atlantic water type to the Mediterranean a unique example of interaction between the atmosphere and the ocean, as demonstrated by the schematic three-dimensional block diagram of the deep currents and the salinity distribution during winter (Fig. 8). In the foreground the Mediterranean Sea is cut open along the axis of the Levantine intermediate current projected along 35°N. Perpendicular to this longitudinal section the sea is also cut open along 6°E and 19°E in order to demonstrate in cross sections the origin of the two main types of deep and bottom waters by the processes of thermal-haline vertical convection (medium arrows). At the surface four types of water are formed which spread out either by measurable currents (thick arrows) or by weaker advective processes (thin, small arrows) in four core layers and operate under mixing, particularly with the intermediate water, the renewal and the ventilation all the way to the bottom of the four great and deep and four smaller and shallower basins of the Mediterranean Sea. The distribution of salinity gives some indications of ascending advection by which the main cycles of vertical circulation are closed.

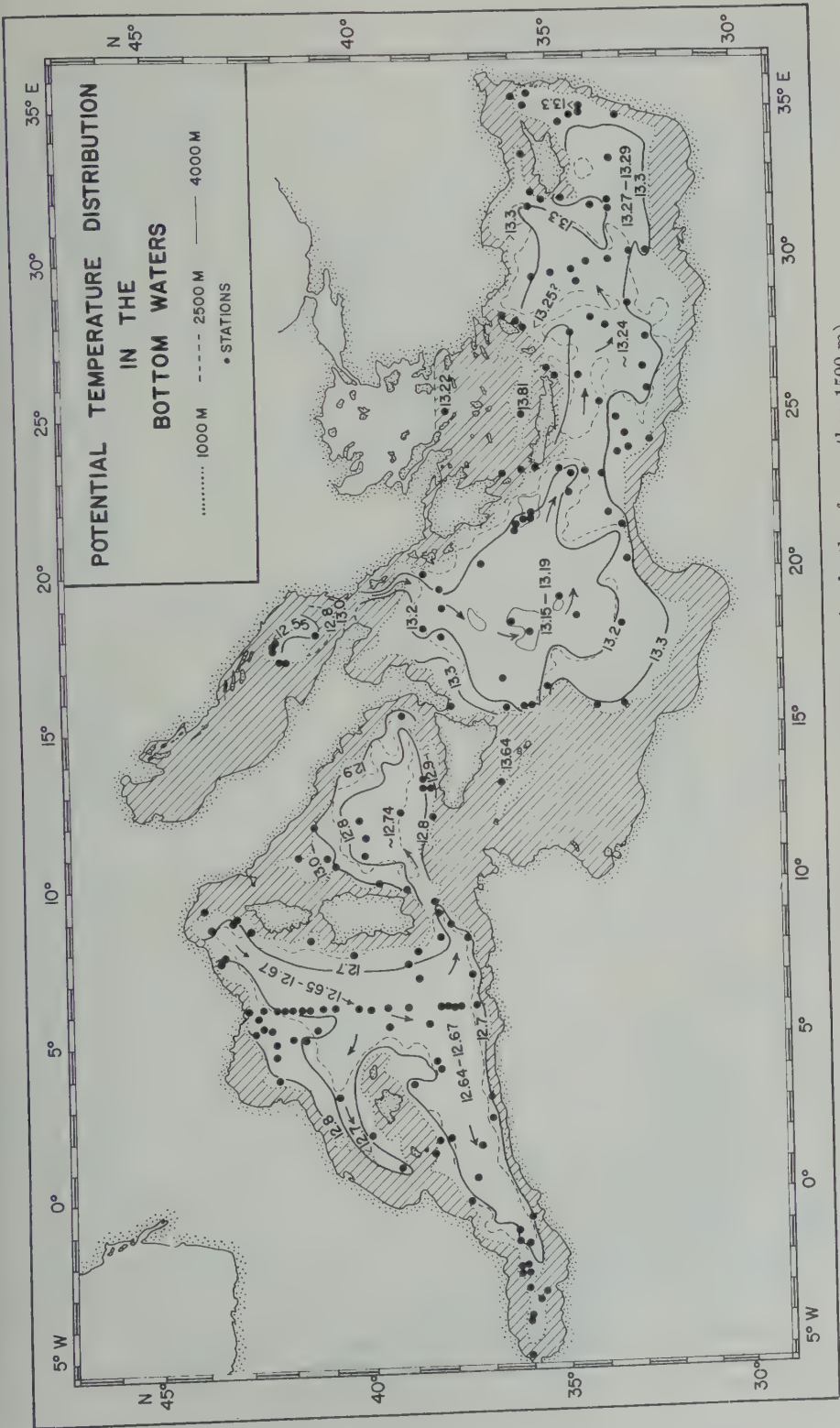


Fig. 7. Map of the potential bottom temperature (at depths of more than 1500 m).

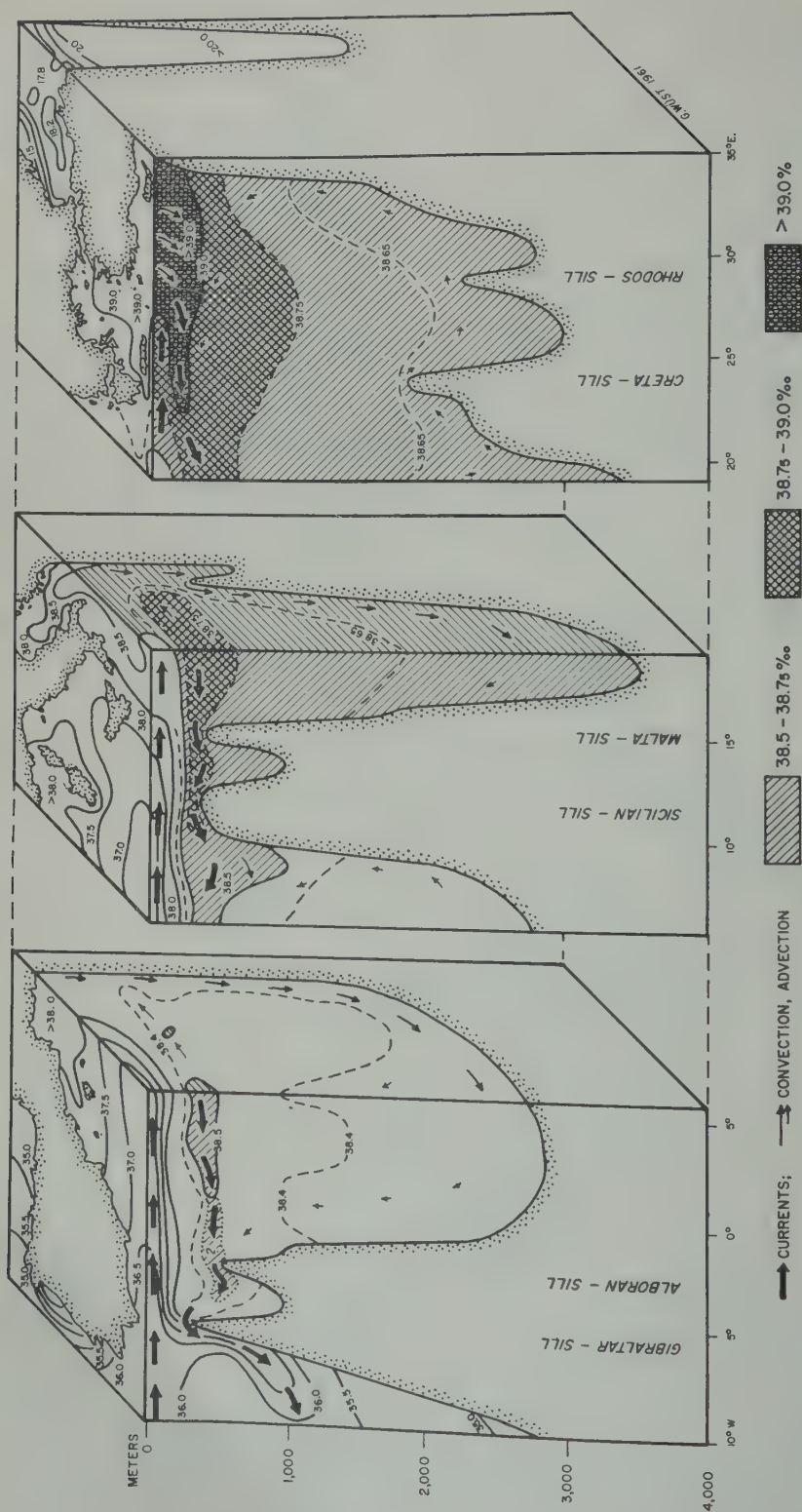


Fig. 8. Schematic block diagram of vertical circulation and distribution of salinity in the Mediterranean Sea during winter.

REFERENCES

- io, A., and L. Trotti, Sur la formation d'eau
nd et d'eau profonde dans la Mer Ligurienne,
ers océanog., 13, 4, Apr. 1961.
- pe, H., and Paul Tschernia, Quelques traits
raux de l'hydrologie Méditerranée, *Cahiers*
og., 12, 8, Sept.-Oct., 1960.
- M. J., The sources of deep water of the
rn Mediterranean Sea, *J. Marine Research*,
s Foundation, 10, 1951.
- Wüst, Georg, Die Tietenzirkulation des Mittel-
ländischen Meeres in den Kernschichten des
Zwischen- und des Tieferwassers, *Deut. Hydro-*
graph. Z., 13(3), Sept. 1960.
- Wüst, Georg, Das Bodenwasser und die Vertikal-
zirkulation des Mittelländischen Meeres, *Deut.*
Hydrograph. Z., 14, in press, 1961.

(Manuscript received May 3, 1961;
revised July 21, 1961.)

General Theory of Dispersion in Porous Media

A. E. SCHEIDEGGER

University of Illinois¹
Urbana, Illinois

Abstract. The possibilities of generalizing the dispersion equations of flow through porous media are investigated. Based on the hypothesis ('Bear's hypothesis') that only that part of each velocity component is of significance which is either parallel or normal to the mean flow direction, the general form of the dispersion is deduced. The dispersivity becomes a tensor of the fourth rank. It has such symmetry properties that it contains only 36 instead of 81 independent components in the general case of an anisotropic porous medium. In isotropic media there are only two dispersivity constants. The latter result had already been deduced by Nikolaevskii. The connection of the dispersivity tensor with a tensor which had previously been constructed by Bear is demonstrated.

Introduction. When a fluid passes through a porous medium, it is customary to describe the process in an entirely macroscopic fashion. Producing the filtration velocity q , i.e. the volume of fluid that passes through a unit of area in unit time, it is possible to treat the flow in a manner which is very much like the treatment of streamline flow in bulk quantities of viscous liquids.

In fact, owing to the complexities of the channels in a porous medium, the passage of a fluid through the medium is a very complicated phenomenon. Owing to 'collisions' with the walls of the pore system, individual elemental volumes of the fluid have a very complicated flow path through the medium which can be treated macroscopically as the path of a random walk. In many applications it is quite sufficient to neglect the microscopic behavior of individual fluid elements as they proceed through the porous medium, but in others this may not be true. The complexity of the pore system causes individual fluid elements to be mixed with each other. Thus, if a tracer be injected at one end of a porous medium, it becomes *dispersed* as it passes through the medium. The term *dispersion* is suggested by the writer [Scheidegger, 1954] to distinguish this process from *diffusion* to which it is superficially similar. As is seen from the above remarks, dispersion is due to the complexities of the pore system; diffusion, on the other hand, is caused by the intrinsic motion of the molecules.

In the first attempts at describing dispersion [Scheidegger, 1954], the diffusion was taken as isotropic. In this case, the motion of a tracer through a porous medium is described by a diffusivity equation with a mass transport term:

$$\partial\psi/\partial t = \text{div } \mathbf{v} (D \text{ grad } \psi) - \mathbf{v} \text{ grad } \psi \quad (1)$$

Here, ψ represents the tracer concentration, D is a factor of dispersion, \mathbf{v} is the macroscopic velocity with which the fluid passes through the porous medium ($\mathbf{v} = \mathbf{q}/P$ if P is the porosity) and t is time. Equation 1 has been experimentally verified for the one-dimensional case.

Theoretical arguments yielded the result that there are two possible connections of the factor of dispersion with the velocity:

$$D = av \quad (2)$$

$$D = av^2 \quad (3)$$

[see Scheidegger, 1957].

Experiments seem to indicate that the first of these forms corresponds to physical reality. This equation implies that there is no mixing of material on adjacent streamlines. For our further arguments, we shall therefore only consider the case in which D is proportional to v . The constant a (with the dimension of length) has been called 'geometrical dispersivity' of the porous medium.

After equation 1 had been set up, experimental evidence showed that D cannot, in fact, be treated as a scalar quantity. In an isotropic porous medium, *longitudinal* as well as *transverse* dispersion occurs, meaning dispersion parallel and transverse to the direction of mean flow.

¹Visiting Professor of Geophysics. Permanent affiliation: Department of Mathematics, University of Alberta, Calgary Campus; Calgary, Alberta, Canada.

The problem is thus posed of generalizing (1) so as to make it applicable, first, to an anisotropic porous medium and, second, to an isotropic porous medium. The latter case, in fact, had already been considered by Nikolaevskii [1959]; we shall confirm Nikolaevskii's result by specifying the constants for a general anisotropic porous medium so that the latter becomes isotropic.

2. *Bear's analysis of variances.* Before proceeding to the solution of the general problem formulated in section 1, we may draw the attention to a remarkable paper by Bear [1961] in which an attempt at a solution, applicable to a restricted case, was made.

Bear did not start with the fundamental equation of motion, but rather with a special solution thereof, representing the motion of a tracer point through the porous medium. If the porous medium be homogeneous, and the mean displacement take place along the x direction only, then the concentration is given by a gaussian function with variance σ_x^2 around the mean position ($x = L$) in the x direction, and σ_y^2 in the y direction (the mean position in the y direction is always, say, zero, since the mean flow takes place in the $+x$ direction only). Bear then postulated that (we write G instead of D in Bear's function to avoid confusion with our notation)

$$\sigma_x = \sqrt{2G_I L} \quad (4)$$

$$\sigma_y = \sqrt{2G_{II} L} \quad (5)$$

which corresponds to the notion of 'geometrical dispersion,' inasmuch as it is assumed that the variances depend on the mean distance traveled only. Bear then set himself the problem of generalizing (4) and (5) so that the generalized equations would be applicable to an isotropic medium in two dimensions in an arbitrary coordinate system. Since the variance $(\sigma^2)_{ik}$ in a gaussian distribution transforms as a tensor, one can state as a working hypothesis

$$(\sigma^2)_{ik} = 2G_{iklm} L_{lm} \quad (6)$$

where the summation convention is understood as being used and

$$L_{lm} = \begin{pmatrix} L \cos^2 \beta & L \sin \beta \cos \beta \\ L \sin \beta \cos \beta & L \sin^2 \beta \end{pmatrix} \quad (7)$$

if β denotes the angle between the direction of L and the l axis. The reason for assuming (6) is

connected with the fact that it is not the component L_i of the displacement L which determines the degree of dispersion, but only the component of L_i in the direction of L and orthogonal to it ('Bear's hypothesis'). The latter components are the components of L_{ik} .

Bear, by calculating the required variances constructed the following tensor G_{iklm} :

$$\left. \begin{aligned} G_{1111} &= G_I \\ G_{1112} &= 0 \\ G_{1121} &= 0 \\ G_{1122} &= G_{II} \\ G_{1211} &= 0 \\ G_{1212} &= \frac{1}{2}G_I - \frac{1}{2}G_{II} \\ G_{1221} &= \frac{1}{2}G_I - \frac{1}{2}G_{II} \\ G_{1222} &= 0 \\ G_{2111} &= 0 \\ G_{2112} &= \frac{1}{2}G_I - \frac{1}{2}G_{II} \\ G_{2121} &= \frac{1}{2}G_I - \frac{1}{2}G_{II} \\ G_{2122} &= 0 \\ G_{2211} &= G_{II} \\ G_{2212} &= 0 \\ G_{2221} &= 0 \\ G_{2222} &= G_I \end{aligned} \right\} \quad (8)$$

Bear did not proceed to establish a connection between G_{iklm} as given in (8) and any equation of motion. As it stands in Bear's paper, the tensor (8) only represents the variance of the tracer distribution (which is zero initially) after a macroscopic displacement by the vector L has taken place. Bear does not show that (8) is the required form for a general isotropic medium, nor does he show that (8) is invariant under an arbitrary rotation of the coordinates. The latter property, however, can be verified easily by direct calculation.

Thus, the task remains to establish the general form of the equation of motion corresponding to (1) for an anisotropic medium. Then this is to be reduced to an isotropic medium, and the most general form for that case is to be written

Finally, it may be shown that (8) represents a special case hereof.

Physical postulates. To set up the most general form of the dispersion equations, we take account of the basic physical aspects of the process.

It is known that during macroscopic flow with velocity v_i in a porous medium, dispersion is anisotropic. The flow equation must be valid in any Cartesian coordinate system, and hence we start from isotropic dispersion in the following form:

$$\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial x_i} D \frac{\partial \psi}{\partial x_i} - v_i \frac{\partial \psi}{\partial x_i} \quad (9)$$

However, experiments show that D is not isotropic, but that transverse dispersion may be neglected. Since we require invariance under coordinate transformations, D must be treated as a tensor. Thus, we write

$$\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial x_i} D_{ik} \frac{\partial \psi}{\partial x_k} - v_i \frac{\partial \psi}{\partial x_i} \quad (10)$$

Experiment also shows that D is a linear function of the velocity. If D were isotropic, one would be tempted to write

$$D = a_i v_i \quad (11)$$

but this is not isotropic. Furthermore, Bear noted that D cannot depend linearly on v_i (Bear took $\int v_i dt$), since it is the components of v_i in the direction of v and orthogonal to it which are important. Thus, we must write

$$D_{ik} = a_{iklm} v_l v_m / |v| \quad (12)$$

This form also retains the linearity in $|v|$. The tensor a_{iklm} may be called the *geometrical dispersivity tensor* of the porous medium. The equation of motion is thus

$$\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial x_i} a_{iklm} \frac{v_l v_m}{|v|} \frac{\partial \psi}{\partial x_k} - v_i \frac{\partial \psi}{\partial x_i} \quad (13)$$

The above form (13) is an outcome of the physical postulates that we have made.

The general form of the equation of motion has also been arrived at by Nikolaevskii [1953], starting from an analogy with the statistical theory of turbulence. Nikolaevskii treated a medium which is *isotropic* and used the theorem of the statistical theory of turbulence which states that the point tensor of dispersion must be a tensor of even rank [Batchelor, 1953]. He thus assumes the form (13) as the simplest

that is consistent with this requirement and pursuing the isotropy condition further, he immediately arrived at the result which we shall deduce at the end of section 5. Nikolaevskii did not consider the anisotropic case.

4. Symmetry properties of the dispersivity tensor. The dispersivity tensor as contained in the equation of motion (13) possesses 81 components.

Fortunately, there are some symmetry properties which reduce the components to 36, even in the general case of a completely anisotropic porous medium.

The first symmetry property is obvious. One has

$$a_{iklm} = a_{ikml} \quad (14)$$

For, if the terms indicated in (14) were not equal, they could be made equal (taking for each term the mean value of the two terms), as this would not affect the equation of motion. One can thus restrict himself to considering only tensors that are symmetric in l and m .

The second symmetry property is

$$a_{iklm} = a_{kilm} \quad (15)$$

This symmetry property is not very obvious as one must take recourse to Onsager's principle of microscopic reversibility in order to justify it. The quantity

$$x_k = \partial \psi / \partial x_k \quad (16)$$

represents a 'force' (the concentration gradient); the quantity

$$j_i = a_{iklm} \frac{v_l v_m}{|v|} \frac{\partial \psi}{\partial x_k} \quad (17)$$

represents a 'flux' (the flow of tracer). The flux-force equations, then, are of the form

$$j_i = L_{ik} X_k \quad (18)$$

and from Onsager's principle it follows that

$$L_{ik} = L_{ki} \quad (19)$$

This leads to (15). The Onsager relation is somewhat naively applied in the above case, but the procedure is entirely analogous to that commonly applied to the heat flow in crystals [Nye, 1957, p. 209].

There seem to be no other symmetry properties of the tensor a_{iklm} in the general case. One is thus left with 36 independent components in the dispersivity tensor. It should be noted that a_{iklm} has more independent components than

for example, the tensor C_{iklm} of the elastic constants of a triclinic crystal. The latter has an additional symmetry property,

$$C_{iklm} = C_{lmik}$$

which reduces the number of independent components to 21. Bear assumed this type of symmetry *a priori* for his tensor G_{iklm} , but he did not give a justification therefor. It is possible that such a symmetry does indeed exist, but the present writer has not been able to find a justification for it. Thus, it seems that one has to reckon with 36 independent dispersivity constants in the general case of an anisotropic porous medium.

It is customary to write a tensor with four indices which possesses only 36 components that are different from zero, in the form of a 6×6 matrix. We shall denote the latter by $a_{\alpha\beta}$ (Greek indices) when α, β run from 1 to 6. The connection between Greek and Latin subscripts is as given in Table 1. In crystals, the corresponding matrix C is symmetric, but a is not symmetric.

5. *Isotropic porous media.* It remains to deduce the general form of the dispersivity tensor for an isotropic porous medium.

We proceed in a fashion similar to that followed when the elastic properties of crystals are deduced. We require numerical symmetry for rotations about the three spatial axes x, y, z by 90° , and we make a final rotation by an arbitrary amount. The formalism cannot be copied from crystals because the matrix $C_{\alpha\beta}$ is symmetric and our matrix $a_{\alpha\beta}$ is not. Nevertheless, the final result turns out to be the same, but we must carry out the calculations anew. Thus, we start with arbitrary values of our 36 dispersivity constants

$$a_{\alpha\beta} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{pmatrix} \tag{20}$$

Making a rotation around the z axis by 90° yields $x \rightarrow y, y \rightarrow -x$. Hence, from Table 1: $1 \rightarrow 2, 2 \rightarrow 1, 3 \rightarrow 3, 4 \rightarrow -4, 5 \rightarrow 6, 6 \rightarrow -5$. Thus, after the rotation the matrix is

TABLE 1. Connection between Greek and Latin Indices

(i, j)	i, j	α
$x \ x$	1 1	1
$y \ y$	2 2	2
$z \ z$	3 3	3
$x \ y$	1 2	4
$x \ z$	1 3	5
$y \ z$	2 3	6

$$a_{\alpha\beta} = \begin{pmatrix} a_{22} & a_{21} & a_{23} & -a_{24} & a_{26} & -a_{25} \\ a_{12} & a_{11} & a_{13} & -a_{14} & a_{16} & -a_{15} \\ a_{32} & a_{31} & a_{33} & -a_{34} & a_{36} & -a_{35} \\ -a_{42} & -a_{41} & -a_{43} & a_{44} & -a_{46} & a_{45} \\ a_{62} & a_{61} & a_{63} & -a_{64} & a_{66} & -a_{65} \\ -a_{52} & -a_{51} & -a_{53} & a_{54} & -a_{56} & a_{55} \end{pmatrix} \tag{21}$$

Requiring that the two matrices must be numerically identical yields

$$\left. \begin{aligned} a_{11} &= a_{22} \\ a_{12} &= a_{21} \\ a_{13} &= a_{23} \\ a_{14} &= -a_{24} \\ a_{15} &= a_{26} = -a_{26} = 0 \\ a_{16} &= -a_{25} = a_{25} = 0 \\ a_{31} &= a_{32} \\ a_{34} &= -a_{34} = 0 \\ a_{35} &= a_{36} = -a_{35} = 0 \\ a_{41} &= -a_{42} \\ a_{43} &= -a_{43} = 0 \\ a_{45} &= -a_{46} = -a_{45} = 0 \\ a_{51} &= a_{62} = -a_{62} = 0 \\ a_{52} &= a_{61} = -a_{61} = 0 \\ a_{53} &= a_{63} = -a_{63} = 0 \\ a_{54} &= -a_{64} = a_{64} = 0 \\ a_{55} &= a_{66} \\ a_{56} &= -a_{65} \end{aligned} \right\} \tag{22}$$

trix, therefore, reduces to

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & 0 & 0 \\ a_{12} & a_{11} & a_{13} & -a_{14} & 0 & 0 \\ a_{31} & a_{31} & a_{33} & 0 & 0 & 0 \\ a_{41} & -a_{41} & 0 & a_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{55} & a_{56} \\ 0 & 0 & 0 & 0 & -a_{56} & a_{55} \end{pmatrix} \quad (23)$$

thus seen that numerical invariance for a rotation around the z axis almost requires that the matrix be symmetric.

We now also require invariance for a 90° rotation around the x axis, we have $z \rightarrow -y$, $y \rightarrow z$, or $1 \rightarrow 1, 2 \rightarrow 3, 3 \rightarrow 2, 4 \rightarrow 5, 5 \rightarrow 4, 6 \rightarrow 6$. Hence, the last matrix becomes

$$\begin{pmatrix} a_{11} & a_{13} & a_{12} & a_{15} & 0 & 0 \\ a_{13} & a_{11} & a_{12} & -a_{15} & 0 & 0 \\ a_{21} & a_{21} & a_{22} & 0 & 0 & 0 \\ a_{51} & -a_{51} & 0 & a_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{44} & a_{46} \\ 0 & 0 & 0 & 0 & -a_{46} & a_{44} \end{pmatrix} \quad (24)$$

adds the following requirements to the previous ones

$$\left. \begin{aligned} a_{11} &= a_{22} = a_{33} \\ a_{12} &= a_{13} \\ a_{14} &= a_{15} = 0 \\ a_{31} &= a_{21} \\ a_{44} &= a_{55} \\ a_{41} &= a_{51} = 0 \\ a_{56} &= a_{46} = 0 \end{aligned} \right\} \quad (25)$$

we have

$$a_{\alpha\beta} = \begin{pmatrix} a_{11} & a_{12} & a_{12} & 0 & 0 & 0 \\ a_{12} & a_{11} & a_{12} & 0 & 0 & 0 \\ a_{12} & a_{12} & a_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{44} \end{pmatrix} \quad (26)$$

The matrix is at last symmetric. It is the dispersivity matrix for a porous medium with cubic symmetry. This corresponds entirely to the elastic matrix for a cubic crystal. To have complete symmetry for the whole group of rotations, we must set

$$a_{44} = \frac{1}{2}(a_{11} - a_{12}) \quad (27)$$

This is the most general dispersivity matrix that is possible for an isotropic porous medium.

In tensor notation we have for a_{iklm} :

$$a_{1111} = a_{2222} = a_{3333} = a_{II} \quad (28)$$

$$\begin{aligned} a_{1122} &= a_{1133} = a_{2233} = a_{2211} \\ &= a_{3311} = a_{III} \end{aligned} \quad (29)$$

$$\begin{aligned} a_{1212} &= a_{1313} = a_{2323} = a_{2121} \\ &= a_{3131} = a_{3232} = a_{1221} = a_{1331} \\ &= a_{2332} = a_{2112} = a_{3113} = a_{3223} \\ &= \frac{1}{2}(a_{II} - a_{III}) \end{aligned} \quad (30)$$

All other terms are zero. The result (28) to (30) had also been obtained by Nikolaevskii [1959].

The physical meaning of a_{II} , a_{III} can clearly be elucidated as follows. We assume that the velocity has the components $(v, 0, 0)$, that the velocity is constant, and that the porous medium is homogeneous. Then the fundamental equation becomes

$$\frac{\partial \psi}{\partial t} = a_{II} v \frac{\partial^2 \psi}{\partial x_1^2} + a_{III} v \frac{\partial^2 \psi}{\partial x_2^2} + a_{III} v \frac{\partial^2 \psi}{\partial x_3^2} \quad (31)$$

It thus turns out that the quantities $a_{II}v$ and $a_{III}v$ are the longitudinal and transverse factors of dispersion as they have been measured:

$$a_{II}v = D_I \quad (32)$$

$$a_{III}v = D_{II} \quad (33)$$

In Bear's investigation these quantities are proportional to his variances. The tensor given by him (2.5) turns out to be a special case of the dispersivity tensor. It may also be concluded that Bear's tensor is the most general one that can be constructed for an isotropic medium in two dimensions.

Acknowledgment. The above investigation was carried out during a term when I was at the University of Illinois. The kind invitation of Professors Read and Rose and their hospitality during my stay at Urbana are gratefully acknowledged.

REFERENCES

- Batchelor, G. K., *Theory of Homogeneous Turbulence*, Cambridge University Press, London, 1953.
- Bear, T., On the tensor form of dispersion in porous media, *J. Geophys. Research*, **66**, 1185-1197, 1961.
- Nikolaevskii, V. N., Konvektivnaya diffuziya v poristykh sredakh, *Prikl. Mat. Mekh.*, **23**, no. 6, 1042-1050, 1959.
- Nye, T. F., *Physical Properties of Crystals*, Clarendon Press, Oxford, 1957.
- Scheidegger, A. E., Statistical hydrodynamics of porous media, *J. Appl. Phys.*, **25**, 994-1000, 1954.
- Scheidegger, A. E., On the theory of flow of miscible phases in porous media, *Compt. rend. Ass. C. Toronto, Assoc. Intern. Hydrol. Scient.*, **2**, 236-240, 1957.

(Manuscript received May 25, 1961.)

Statistical Geometry of Porous Media

H. D. FARA¹ AND A. E. SCHEIDEGGER²

*College of Engineering
University of Illinois, Urbana, Illinois*

Abstract. The geometrical characterization of porous media is an important task in problems occurring in hydrology. The usual geometrical quantities, such as porosity, specific surface, etc., are not sufficient to describe a given porous medium fully, and a different approach must therefore be sought. An attempt is made to characterize a porous medium from data which can be read off a photomicrograph. Four possibilities are considered: a correlation function, a spectral analysis in terms of harmonic functions, a spectral analysis in terms of other orthogonal functions, and a spectral analysis of a specially constructed characteristic function of the porous medium in terms of harmonic functions. The two types of harmonic analyses yield the most satisfactory descriptions of a porous medium. The method is applied to an actual example.

Introduction. Porous media are important in a variety of contexts. They are encountered, for instance, in all branches of geophysics in which the flow of fluids within the earth's crust is of interest. In general, macroscopic empirical laws are used to describe the flow of fluids through porous media which, in many practical applications, yield completely satisfactory results [see, e.g., Scheidegger, 1960]. However, it is desirable to have a description of the geometrical properties of a porous medium in somewhat greater detail. Geometrically, a porous medium is defined by giving the analytical equation of the surface which bounds the pore space. For any practical purpose this is impossible to accomplish. Average geometrical quantities, such as porosity (denoted here by ϕ) and specific surface (ratio of internal surface to bulk volume), are conceptually acceptable (and can be measured), but this is not true for the concept of 'pore size.' Since the pores form an extremely complicated system of hollows within a solid body, it is not easy to determine a pore size in a conceptually satisfactory fashion.

Nevertheless, there is a great need for establishing a better geometrical characterization of a porous medium than is afforded by the mere determination of porosity and specific surface. It is particularly desirable to have a criterion

by which it can be decided by looking at two porous media through a microscope whether or not they are identical in a statistical sense.

Some methods are proposed by which a statistical characterization of the geometry of a porous medium might be accomplished by specifying a set of numbers or a function.

2. *Statistical description of a porous medium.* From the foregoing remarks it is obvious that in order to characterize the geometry of a porous medium, one must somehow take recourse to statistics. Notions such as porosity and specific surface area are essentially of a statistical nature, and it would be desirable to generalize these notions in such a fashion as to obtain a more complete description of the geometry of a porous medium than is represented by these two terms alone.

To arrive at such a statistical description, let us proceed as follows. Let us assume that an arbitrary line be drawn through a given porous medium whose geometry is to be described. Points on the line are to be defined by giving their arc length s from an arbitrarily chosen origin. Then, for certain values of s the line will pass through void spaces; for other values of s it will pass through filled spaces. We then introduce a function $f(s)$ of the arc length s of the line which is defined as follows: the value of f is defined as $+1$ if the line at s passes through void space; it is defined as equal to -1 if the line at s passes through filled space (see Fig. 1).

After the definition of the function $f(s)$ we can start to do some statistical analyzing. First,

¹Research Associate.
²Visiting Professor of Geophysics. Permanent address: Department of Mathematics, University of Alberta, Calgary, Alberta, Canada.

we can define the mean of f :

$$\bar{f} = \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} f(s) ds}{\int_{-s}^{+s} ds} \quad (1)$$

If the porous medium is homogeneous and isotropic (in a statistical sense), then the value of \bar{f} does not depend on the direction or position of the line. Of course, as in all considerations of probability theory, those lines must be excluded that have been drawn in some special way. In fact, $f(s)$ ought to be treated as a random function of which the particular $f(s)$ found in a special case is a *realization*. The mathematical theory of such random functions has been summarized, e.g. in a paper by Moyal [1949] in which it is shown that the calculus of random functions is very much the same as that of ordinary functions, provided suitable generalizations in the definition of integration, convergence, etc., are adopted.

The mean value \bar{f} of f can be expressed in terms of the porosity P of the medium:

$$\bar{f} = 2P - 1 \quad (2)$$

The higher moments of our function (i.e. \bar{f}^2, \bar{f}^3) are equal to either \bar{f} (odd moments) or 1. This can easily be demonstrated as follows: $f^2(s)$ is always 1. Hence, if n be even, we have

$$\begin{aligned} \bar{f}^n &= \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} f^n(s) ds}{\int_{-s}^{+s} ds} = \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} (f^2)^{n/2} ds}{\int_{-s}^{+s} ds} \\ &= \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} 1^{n/2} ds}{\int_{-s}^{+s} ds} = 1 \end{aligned} \quad (3)$$

and if n be odd

$$\begin{aligned} \bar{f}^n &= \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} f^n(s) ds}{\int_{-s}^{+s} ds} = \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} f(f^2)^{n-1/2} ds}{\int_{-s}^{+s} ds} \\ &= \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} f \cdot 1^{n-1/2} ds}{\int_{-s}^{+s} ds} = \bar{f} \end{aligned} \quad (4)$$

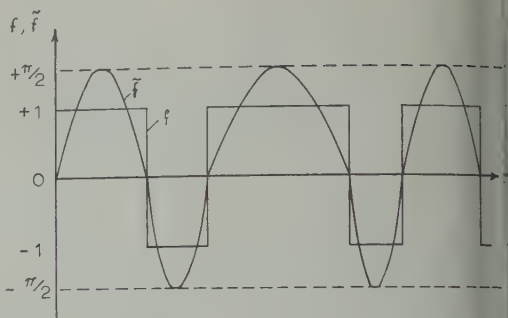


Fig. 1. Graph of $f(s)$ and $\bar{f}(s)$ in a general case.

Thus, the analysis of moments does not lead us any further.

In order to proceed with our analysis, it is inconvenient to have a function f whose mean is not zero. We therefore define a new function f' which is connected with the old function f as follows:

$$f'(s) = f(s) - (2P - 1) \quad (5)$$

The mean of f' is zero.

A very useful statistical quantity based upon $f(s)$ is the autocorrelation function $R(\Delta)$:

$$R(\Delta) = \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} f(s)f(s + \Delta) ds}{\int_{-s}^{+s} ds} \quad (6)$$

We know that $R(0) = 1$ (see eq. 3). The autocorrelation function will not tend to zero for $\Delta \rightarrow \infty$ because $f(s)$ is not of zero mean. It tends toward the square of its mean, i.e.

$$R(\infty) = \bar{f}^2 = (2P - 1)^2 \quad (7)$$

The autocorrelation function affords a statistical description of the porous medium.

Again, it is useful to introduce f' (of zero mean) instead of f . We then have

$$R'(\Delta) = \lim_{s \rightarrow \infty} \frac{\int_{-s}^{+s} f'(s)f'(s + \Delta) ds}{\int_{-s}^{+s} ds} \quad (8)$$

Particular values of the function R' are

$$R'(0) = 4(P - P^2) \quad (9)$$

$$R'(\infty) = 0 \quad (10)$$

statistical description of a porous medium obtained by giving the function $R(\Delta)$).

Spectral theory of a porous medium. The character of the function $f(s)$, with its random character from 1 to -1 , is that of a *random function*. Random functions are very well known in mathematical physics from the investigation of Brownian motion. As noted above, an excellent summary of the calculus of random functions has been given, for instance, by Moyal [1949].

If the porous medium is homogeneous, the random function $f(s)$ must be a *stationary* random function, implying that all the joint probability distributions (taken at s_1, s_2, \dots, s_n) must be the same. The random functions $f(s)$ are independent of the choice of the origin of s . Under these circumstances the random function $f(s)$ has a *spectrum*, and we can therefore write

$$f(s) = \int_{-\infty}^{+\infty} e^{iks} dZ(k) \quad (11)$$

is an orthogonal process. Every stationary random process possesses a Fourier expansion of the above type. This fact is connected with the condition that $f(s)$ is expressible by an operator T_s acting upon $f(s_0)$ so that

$$f(s) = T_s f(s_0) \quad (12)$$

$T_s T_t = T_{s+t}$. The only possible operator having this property is

$$T_s = \int_{-\infty}^{+\infty} e^{iks} dE_k \quad (13)$$

E_k is a resolution of the identity operator. We always have the possibility of writing

$$\begin{aligned} T_s f(s_0) &= \int_{-\infty}^{+\infty} e^{iks} d(E_k f(s_0)) \\ &= \int_{-\infty}^{+\infty} e^{iks} dZ(k) \end{aligned} \quad (14)$$

is indeed equation 11. The inverse of the spectral representation is

$$dZ(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1 - e^{-iks}}{is} f(s) ds \quad (15)$$

The formulas assume a somewhat more familiar look if it be assumed that a spectral density exists:

$$dZ(k) = \varphi(k) dk \quad (16)$$

Such a spectral density does not exist for $f(s)$, but it may exist for $f'(s)$, for example, which is the function defined in (5). Then

$$f'(s) = \int_{-\infty}^{+\infty} e^{iks} \varphi(k) dk \quad (17)$$

$$\varphi(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iks} f'(s) ds \quad (18)$$

If real and imaginary terms are separated,

$$\varphi(k) = a(k) + ib(k) \quad (19)$$

$$a(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \cos ks f'(s) ds \quad (20)$$

$$b(k) = -\frac{1}{2\pi} \int_{-\infty}^{+\infty} \sin ks f'(s) ds \quad (21)$$

$$\begin{aligned} f'(s) &= 2 \int_0^{\infty} [a(k) \cos ks \\ &\quad - b(k) \sin ks] dk \end{aligned} \quad (22)$$

An inspection of the above formulas shows that changing the origin of s effects a multiplication of $dZ(k)$ by a complex number of unit magnitude. Hence, in order to average over the possible positions of the origin, one should consider the invariant

$$c^2(k) = a^2(k) + b^2(k) = \varphi(k) \varphi^*(k) \quad (23)$$

the star denoting the complex conjugate.

The specification of $c^2(k)$ as a function of k constitutes a reasonable statistical representation of the geometry of a porous medium. Moreover, this representation is independent of the choice of the origin. One may also note that k has the dimension (length) $^{-1}$. The corresponding wavelength

$$\lambda = 2\pi/k \quad (24)$$

can be regarded as a wavelength in the porous medium.

4. *The concepts of homogeneity and isotropy.* Thus far, in the definition of the autocorrelation function and of the spectrum, only superficial reference has been made to the concepts of homogeneity and of isotropy.

The concept of homogeneity is implied in the definitions of $R(\Delta)$ and of $c(k)$ as given above, since, in order to have the same values for these expressions for all possible lines, the medium must be homogeneous in a statistical sense.

Actually, it will be impossible to have the same values for *all* lines, but the probability of encountering a line for which these values differ substantially from those for the majority of the lines must be ‘almost certainly zero’ (for a definition of such ‘almost certain’ probabilities, see *Moyal* [1949]). A porous medium having these properties is necessarily homogeneous and isotropic.

Let us now assume that there are preferred directions in a porous medium. We assume that the lines along which we define our function $f(s)$, and consequently $f'(s)$, are straight and parallel to one direction. We say that the porous medium is *anisotropic*, if the functions $c(k)$ and $R(\Delta)$ depend on the *direction* which is chosen for the line. This can be expressed by redefining the arguments of these functions as *vectors* \mathbf{k} and Δ .

There is no reason to believe that ‘principal’ directions exist in the sense that it would suffice to give the values of the functions in question along these directions only. The functions c and R are true functions of space.

5. *Different orthogonal expansion.* The function $f(s)$ characterizing the porous medium has many discontinuities built into it. Therefore, the thought might occur to use, for the spectral expansion, an orthogonal system of functions which has the discontinuities already inherent in it instead of the system of harmonic functions which consists of continuous functions.

A possible method of doing this will be outlined below; however, when applying this method, an apparently unremovable phase angle appears.

Thus, instead of using the system of trigonometric functions, we wish to see whether a suitable system could be built up from the functions $\text{sgn} \sin ks$ and $\text{sgn} \cos ks$. The symbol sgn denotes the signum function, which is defined as follows:

$$\left. \begin{aligned} \text{sgn } x &= 1 && \text{for } x > 0 \\ \text{sgn } x &= -1 && \text{for } x < 0 \\ \text{sgn } x &= 0 && \text{for } x = 0 \end{aligned} \right\} \quad (25)$$

The set of functions introduced above do not form an orthogonal system. We can see how an orthogonal set of discrete functions can be constructed from a set of functions that is not

orthogonal. Thus, let us consider lines of length L only upon which $f(s)$ (and therewith $f'(s)$) is defined. We then define the following functions

$$\psi_0 = 1/\sqrt{2L} \quad (26)$$

$$\psi_n^{(a)} = (1/\sqrt{2L}) \text{sgn} \cos \pi ns/L \quad (27)$$

$$\psi_n^{(b)} = (1/\sqrt{2L}) \text{sgn} \sin \pi ns/L \quad (28)$$

It is clear that all the functions $\psi^{(a)}$ are orthogonal to all the functions $\psi^{(b)}$, merely because one set is symmetric with regard to the origin and the other is antisymmetric. Hence each system $\psi^{(a)}$ or $\psi^{(b)}$ is to be made orthogonal within itself. Thus, we introduce new functions $\varphi_n^{(a)}$ and $\varphi_n^{(b)}$ which form an orthogonal system. The procedure for accomplishing this is well known; for the b functions, say, it is as follows:

$$\left. \begin{aligned} \varphi_1^{(b)} &= \psi_1^{(b)} \\ \varphi_2^{(b)} &= \mu_{21} \psi_1^{(b)} + \mu_{22} \psi_2^{(b)} \\ \varphi_3^{(b)} &= \mu_{31} \psi_1^{(b)} + \mu_{32} \psi_2^{(b)} + \lambda_{33} \psi_3^{(b)} \quad \text{etc.} \end{aligned} \right\} \quad (29)$$

The equations for finding the μ 's are obtained by expressing the condition that each new φ must be orthogonal to all previous ones and must be normalized. For the first three functions we obtain

$$\left. \begin{aligned} \varphi_1^{(b)} &= \psi_1^{(b)} \\ \varphi_2^{(b)} &= \psi_2^{(b)} \\ \varphi_3^{(b)} &= -\frac{1}{2\sqrt{2}} \psi_1^{(b)} + \frac{3}{2\sqrt{2}} \psi_3^{(b)} \quad \text{etc.} \end{aligned} \right\} \quad (30)$$

A similar procedure can be set up for the a functions.

A ‘spectral’ analysis (in terms of the above orthogonal functions) can now be made of the function f describing a porous medium:

$$a_n = \int_{-L}^{+L} \varphi_n^{(a)} f(s) \, ds \quad (31)$$

$$b_n = \int_{-L}^{+L} \varphi_n^{(b)} f(s) \, ds \quad (32)$$

The coefficients defined by the last two equations afford a ‘spectral’ description of the porous medium. To make these coefficients independent of the length L of the line it is sensible to divide



3. 2. Photomicrograph of porous medium (stone) 92X. Courtesy of Humble Oil and Refining Company. Porosity 26.3 per cent. The line is 17 cm in length.

by $(2L)^{1/2}$; thus

$$a_n'' = a_n/\sqrt{2L} \tag{33}$$

$$b_n'' = b_n/\sqrt{2L} \tag{34}$$

the a_n'' , b_n'' are now the relative spectral densities for the wavelength

$$\lambda = 2L/n \tag{35}$$

Unfortunately, it is now no longer feasible to use the combination

$$c_n''^2 = a_n''^2 + b_n''^2 \tag{36}$$

to get rid of the phases, because (36) is not an invariant of translation. The present description of a porous medium thus has a phase angle built into it which does not render it as satisfactory as a description in terms of sines and cosines.

Rounding off the corners. A spectral analysis of a porous medium may also be made by attempting to modify the function $f(s)$ which characterizes the medium in such a fashion that it is no longer necessary to deal with discontinuities.

TABLE 1. Spectral Coefficients of a Porous Medium

n	Function f , c_n	Function \tilde{f} , c_n
0	0.464	0.465
1	0.267	0.268
2	0.606	0.714
3	0.272	0.457
4	0.154	0.052
5	0.277	0.289
6	0.232	0.498
7	0.244	0.474
8	0.439	0.478
9	0.260	0.247
10	0.353	0.410
11	0.142	0.110
12	0.158	0.154
13	0.229	0.278
14	0.088	0.085

The question comes to mind, therefore, whether it would not be possible to round off the corners of $f(s)$ in some consistent fashion. The rounding off has to be accomplished in such a way that the essential features of the porous medium are preserved.

We suggest that in $f(s)$ each 'square wave' be replaced by a sine wave of the same area and the same half-wavelength. This will produce a discontinuity of the slope at the points where the zero line is crossed, but the sharp corners will be removed. The new function may be denoted by $\tilde{f}(s)$ (see Fig. 1).

Each square wave is thus replaced by a sine wave. The amplitude of the sine waves must be chosen as $\pi/2$ so that the equal area requirement is satisfied. The new function $\tilde{f}(s)$ can be dealt with in exactly the same manner as the old function $f(s)$. In particular, the mean of $\tilde{f}(s)$ is the same mean as that of $f(s)$ and is connected with the porosity of the porous medium by equation 2. In accordance with previous procedures, a function of zero mean, $\tilde{f}'(s)$, which has the same relation to $\tilde{f}(s)$ as $f'(s)$ has to $f(s)$, can again be introduced, making possible exactly the same spectral analysis of $\tilde{f}(s)$ as has been suggested in the preceding sections for $f(s)$.

7. *Example.* The spectral analysis suggested in sections 3 and 6 has been performed for an actual porous medium. Figure 2 is a photo-

micrograph of a rock which had been impregnated with Wood's metal and had then been cut and polished.

In the formulas given in sections 3 and 6 the spectral analysis was made for a line of infinite length. In practice, a line will always have a finite length L , and the spectral representation is as follows:

$$f(s) = \sum_{n=-\infty}^{+\infty} a_n e^{i2\pi ns/L} \tag{37}$$

where

$$a_n = a_{-n}^* \tag{38}$$

Of interest is only

$$c_n = + \sqrt{a_n a_{-n}} \tag{39}$$

since the a 's contain a phase angle. In our example, the length of the line in Figure 2 is 0.17 cm.

The University of Illinois Illiac digital computer was programmed to give the spectral coefficients c_n for the functions $f(s)$ and $\bar{f}(s)$ directly from the input $f(s)$. The results of the calculations are shown in Table 1. The first coefficient ($n = 0$) is connected with the porosity P of the line by the equation

$$c_0 = |2P - 1| \tag{40}$$

In the present example the porosity P along the line comes out as 26.8 per cent, which is in good agreement with the volumetric porosity of the sandstone mentioned above.

An inspection of the remaining coefficients shows that in both cases the prevalent wavelength is $L/2$, which is representative of the two large filled spaces at the ends of the line which together take up about one-fourth of its length. The next significant wavelength is around L which represents grains of diameter $L/16$.

The above discussion shows that the suggested spectral coefficients do indeed characterize a given porous medium.

Acknowledgments. The present study was carried out during the stay of one of us (A.E.S.) as Visiting Professor of Geophysics at the University of Illinois. The efforts of Professors Read and Rott in making this visit possible are gratefully acknowledged. We also wish to express our gratitude to the University of Illinois for making its Illiac computer available and to the Humble Oil and Refining company for supplying the photomicrographs.

REFERENCES

Moyal, J. E., Stochastic processes and statistical physics, *J. Roy. Statist. Soc., B9*, 150-210, 1949.
Scheidegger, A. E., *The Physics of Flow Through Porous Media*, 2nd ed., University of Toronto Press, 1960.

(Manuscript received July 24, 1961.)

Effect of Interstation Correlation on Regression Analysis

NICHOLAS C. MATALAS AND MANUEL A. BENSON

U. S. Geological Survey, Washington 25, D. C.

Abstract. The question is raised as to the reliability of regression relations in which stream flow, the dependent variable, is related to characteristics of the drainage basin. The values of such a dependent variable are often interrelated and thus fail to satisfy one of the basic assumptions of regression theory. The consequences of such conditions are first examined on hydrologic grounds. A statistical demonstration is then presented for a simple regression model in which the dependent variable is intercorrelated. Exact relations are given for the reliability of the regression constant, a , the slope of the regression line, b , and the predicted value of the dependent variable, \hat{y} . Although interstation correlation does not affect the expected values of a , b , and \hat{y} , it does affect their variance. Hence, in order that valid inferences may be made from a regression analysis, the interstation correlation must be taken into consideration when it is present. The results are general and apply to many hydrologic relations, as well as to any other phenomena with similar conditions of interrelation in the dependent variable.

INTRODUCTION

Hydrologists are often interested in studying a hydrologic variable such as river discharge with the physical parameters describing the drainage area. One may wish to know how mean annual floods of the streams in a given area vary with the size of the drainage areas. This variation is usually studied by assembling data for the several rivers that have gaging stations on records and using regression theory to determine a relationship between discharge as the dependent variable and the physical parameters as the independent variables.

In this kind of study it is assumed, among other things which will be discussed, that the values of the dependent variable are uncorrelated. However, hydrologists have long been aware that this assumption is not correct. In a given region, rivers fluctuate more or less in response to a common set of meteorologic factors. The rivers rise in response to a rain storm that more or less affects all the rivers in the region; at another time, the rivers are at low levels owing to a general absence of rain. Thus, the flows of different streams are correlated, and each river therefore repeats some part of the history of its neighbors.

The lack of independence may exist even in regions where rain occurs as localized thunderstorms. For example, correlation may be quite high between annual peak discharges of different streams having but few annual peaks

of the same date because the timing and magnitude of thunderstorm activity may be related over regions that are much larger than the extent of the individual thunderstorms.

The effect of interstation correlation may be examined for a hypothetical region with perfect correlation in the dependent variable (annual peak discharge).

1. Assume 100 stations within a region of widely varying physical characteristics.
2. Suppose that in each year there is only one rainstorm of any consequence in this area. The storm has a uniform intensity over the entire region and lasts exactly 24 hours.
3. Suppose that this storm produces a peak discharge of the same probability; for example, a 10-year peak at each station. This is unlikely, but the assumption is made at this time as a starting point for speculation.
4. Suppose that there is a long-term or basic distribution of floods within the region.
5. Assume a 50-year record over this area with a representative distribution of floods; that is, both the rainfalls and the resulting peak discharges would define frequency curves with little or no scatter of individual points.

What is known about the so-called 50-year flood level (the .02-probability flood)? There is only one experience of it in general over the area. For this reason the 50-year flood in general or the 50-year peak discharge at any station is not known with any confidence. What has been

called the 50-year flood level may actually be (if there is a long-time 'actual') the 80-year level or the 25-year level. The 10-year flood level has been based on five experiences and therefore is better known than the 50-year level; yet a statistic based on five items has an extremely wide confidence band.

Suppose that the 50-year peak discharges (dependent variables) at the 100 stations are correlated with the physical characteristics of the basins (independent variables) and that a multiple-regression equation is developed. What is the significance of such an equation and what confidence may we place in it? The equation is based on 100 sets of data, one set at each station. The dependent variables are all known to be at the same flood level, but it is not known with any confidence what that level may be. Yet the relationship between peak discharges and physical characteristics for that level of flood has been determined with some exactness, based on 100 sets of data and nearly 100 degrees of freedom. We arrive at a fairly dependable relation between known quantities, the physical characteristics, and peak discharges whose magnitudes are known but whose true recurrence interval is not known. The uncertainty of the recurrence interval does not diminish the confidence limits of the relation with the physical characteristics—on the other hand, the reliability of the regression equation does not add to the knowledge of the true recurrence interval of the discharges.

Of what use, then, is such an equation? Because of the shortness of the records we are faced with the fact that our '50-year flood' is only nominally such. Yet this figure, artificial though it may be, must serve as the basis for our problems of engineering design. There is almost a necessity for acting as though the available period of record has defined the true distribution of floods, though we know this is not so. For our design flood, whether a 50-year flood or not, the regression equation tells us how the so-called 50-year peak discharges vary on each drainage basin.

We have postulated a 50-year record with a given distribution of floods, on the basis of which, for design purposes, we may set values on the 50-year, 25-year, 10-year, etc., peaks. For each of these flood levels we may determine regression equations relating the T -year peak discharge,

Q_T , to the hydrologic characteristics. The relations are commonly linear in terms of logarithms of the original variables:

$$\log Q_T = \log a + b \log B + c \log C + d \log D + \dots$$

and can be expressed more simply as

$$Q_T = aB^b C^c D^d \dots$$

where B, C, D, \dots are the independent variables characterizing the hydrologic conditions and a, b, c, d, \dots are constants of the regression equation.

Suppose now that we experience a second 50-year period during which the annual peak average twice the magnitude of the peaks in the first period. From data of this period we could again set values to the 50-, 25-, and 10-year peaks and again develop regression equations relating them to the hydrologic characteristics. The resulting equations could then be expected to have a values approximately twice those of the first period; the remaining regression constants b, c, d, \dots would be the same as those defined for the first period.

What is the relation of the actual conditions to our model? In regions where snowmelt peaks are common there may be only one peak each year, but in other regions there are generally more than one. Annual peaks at all stations would not be caused by the same storm. The annual peak discharges of each year do not have the same recurrence intervals at all stations. Instead the 25-year peaks, for example, are determined from the frequency curves at each station. Yet the list of 25-year peaks that are used to set up a regression equation defining the 25-year peak in terms of basin and climatic characteristics is based on the experience during a given period with a given level of flood experience. It seems clear that, in the regression equation that is developed, the a constant is dependent on the general level of flood activity and the other regression constants are based on the relation between that level of activity and the hydrologic characteristics.

It appears worth while to investigate more closely the effect of interstation correlation upon the reliability of a relation that may be defined between river discharges and pertinent physical parameters of the catchment areas contributing to each of the rivers.

RELATION AND REGRESSION THEORY IN
HYDROLOGY

$\{y_i\}$ denote the sequence of discharges for the i th gaged drainage area, and let y_i denote the classified (mean, percentile, etc.) value of y from the sequence, where $i = 1, 2, \dots, k$. If x_1, \dots, x_n denote the magnitudes of the physical parameters describing the i th drainage area, the regression of y on $x_1 \dots x_n$ is given by

$$b_0 + b_1(x_1 - \bar{x}_1) + \dots + b_n(x_n - \bar{x}_n) \quad (1)$$

where b_0 is the regression constant and $b_1 \dots b_n$ are the regression coefficients. More commonly, hydrologic relations are such that y and x denote logarithms of discharge and basin characteristics, respectively.

The assumptions which underlie the regression model (equation 1) are as follows: (1) The independent variables, $x_1 \dots x_n$, represent fixed values and hence do not have probability distributions. (2) The residuals from the line of regression are normally distributed. (3) The variance about the line of regression is constant. Hence the values of the dependent variable y are normally independent. Under these assumptions, regression can be treated by the classical methods used in regression analysis. The successful application of regression in hydrology is dependent upon how well the underlying assumptions are satisfied.

The first assumption can be met by pre-selecting the magnitudes of $x_1 \dots x_n$ and then selecting those drainage areas possessing those values. In satisfying this assumption many difficulties would be encountered. However, for predictive purposes, it is best to consider the independent variables as fixed values regardless of how they are selected.

In most hydrologic problems the number of drainage areas in a region is small, so that it is difficult to prove whether the residuals are normally distributed. If the residuals are found to be non-normally distributed, it is necessary to transform the dependent variable in order to normalize the residuals. Transformations such as logarithms, square roots, and cube roots often normalize the residuals, at least within the limits of random sampling errors. The third assumption is not always satisfied in hydrologic data. It frequently happens that

the dispersion about the line of regression increases with an increase in the magnitudes of the dependent and independent variables. The dispersion often can be stabilized by means of transformations. Generally, the transformation used to stabilize the dispersion also normalizes the residuals.

The fourth assumption is more troublesome. Within a given region, the sequence of flows, or some transformation of the flows, for one drainage area varies linearly with the sequence of flows for another drainage area. Hence, the classified values of discharge selected from each of the sequences are correlated and are not mutually independent.

From the above discussion it is seen that the first three assumptions are met or can be met in hydrology. If the classified values of flow are linearly dependent, hence correlated, the fourth assumption is not valid, and the regression analysis must be modified.

To gain an insight into the effect of interstation correlation, we must restrict the regression model to only one independent variable. Thus,

$$Y = a + b(x - \bar{x}) \quad (2)$$

where a is the regression constant and b the regression coefficient. The true regression of y on x is given by

$$y = \alpha + \beta(x - \bar{x}) \quad (3)$$

The coefficients a and b are estimates of α and β , respectively.

If y_i and y_j denote the classified values of flow from the i th and j th drainage areas, respectively, the interstation correlation may be expressed as

$$\text{covariance}(y_i y_j) = \rho_{ij} \sigma_i \sigma_j \quad (4)$$

where ρ_{ij} is the measure of interstation correlation between y_i and y_j , σ_i and σ_j denote the standard deviations of y_i and y_j , respectively. If there are k gaged drainage areas in the region, then there are $k(k-1)/2$ possible interstation correlations.

To determine the effect of interstation correlation on the regression model given by (2) we make the following assumptions: (1) The x 's are fixed variates. (2) The residuals from the line of regression are normally distributed. (3) The variance about the line of regression is a con-

stant. (4) $\sigma_i = \sigma_j$ for all values of i and j . Under these assumptions, the effect of interstation correlation is assessed by determining the following statistics and comparing the values for these statistics with the values for these statistics when there is no interstation correlation: (1) The expectation of a , $E(a)$. (2) The variance of a , $V(a)$. (3) The expectation of b , $E(b)$. (4) The variance of b , $V(b)$. (5) The expectation of an estimated value of y , $E(\hat{y})$. (6) The variance of an estimated value of y , $V(\hat{y})$.

MATHEMATICAL DEVELOPMENT

Regression constant. From (2) and (3),

$$Y_i - y_i = e_i \\ = (a - \alpha) + (b - \beta)(x_i - \bar{x}) \quad (5)$$

where e_i is the residual or difference between the estimated value Y_i and the observed value y_i . The mean value of e may be expressed as

$$\bar{e} = \frac{1}{k} \sum_{i=1}^k e_i = (a - \alpha) \quad (6)$$

Since e is assumed to be normally distributed with zero mean, the expectation of \bar{e} is zero, so that

$$E(a) = \alpha \quad (7)$$

whereby a is an unbiased estimator of α .

The variance of a may be expressed as

$$V(a) = E\left[\frac{1}{k} \sum_{i=1}^k e_i\right]^2 = E(a - \alpha)^2 \quad (8)$$

whereby, on expansion,

$$V(a) = \frac{1}{k^2} E\left[\sum_{i=1}^k e_i^2 + 2 \sum_{j=i+1}^k \sum_{i=1}^{j-1} e_i e_j\right] \quad (9)$$

It has been assumed that the variance of y about the line of regression is a constant. Hence,

$$E(e_i^2) = E(e_j^2) = \sigma_e^2 \quad (10)$$

for all values of i and j , where σ_e^2 is the standard error of estimate, as ordinarily computed. The expectation of $e_i e_j$ is defined as

$$E(e_i e_j) = R_{ij} \sigma_e^2 \quad (11)$$

where R_{ij} denotes the correlation between e_i and e_j and is equal to the correlation between y_i and y_j , ρ_{ij} . Hence, the variance of a can be expressed as

$$V(a) = (\sigma_e^2/k)[1 + (k-1)\bar{R}] \quad (12)$$

where \bar{R} is the average of all possible values of R_{ij} for $i \neq j$.

If $R_{ij} = 0$ for all values of $i \neq j$, (12) reduces to

$$V(a) = \sigma_e^2/k \quad (13)$$

Since $R_{ij} \geq 0$ for all values of $i \neq j$, the interstation correlation makes the variance of a greater than would be the case if there were no interstation correlation. As k tends to infinity, it is seen from (12) that the variance of a tends to

$$V(a) = \bar{R} \sigma_e^2 \quad (14)$$

If $\bar{R} = 0$ (and if k tends to infinity), $V(a) = 0$. In Figure 1 a family of curves is shown as a function of \bar{R} , with $V(a)/\sigma_e^2$ as the ordinate and k as the abscissa.

Regression coefficients. The regression coefficient is defined in the theory of least squares as

$$b = \frac{\sum_{i=1}^k (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^k (x_i - \bar{x})^2} = \frac{\sum_{i=1}^k y_i (x_i - \bar{x})}{\sum_{i=1}^k (x_i - \bar{x})^2} \quad (15)$$

The expected value of y_i is $\alpha + \beta(x_i - \bar{x}) + e_i$, whereby (15) can be expressed as

$$b = \beta + \frac{\sum_{i=1}^k e_i (x_i - \bar{x})}{\sum_{i=1}^k (x_i - \bar{x})^2} \quad (16)$$

Since e_i is independent of $(x_i - \bar{x})$, the expectation of $e_i(x_i - \bar{x})$ equals zero, so that

$$E(b) = \beta \quad (17)$$

Thus, b is an unbiased estimator of β .

The variance of b may be expressed as

$$V(b) = E(b - \beta)^2 = E\left[\frac{\sum_{i=1}^k e_i (x_i - \bar{x})}{\sum_{i=1}^k (x_i - \bar{x})^2}\right]^2 \quad (18)$$

whereby, on expansion,

$$V(b) = \left[\frac{1}{\sum_{i=1}^k (x_i - \bar{x})^2}\right]^2 \cdot E\left[\sum_{i=1}^k e_i^2 (x_i - \bar{x})^2 + 2 \sum_{j=i+1}^k \sum_{i=1}^{j-1} e_i e_j (x_i - \bar{x})(x_j - \bar{x})\right] \quad (19)$$

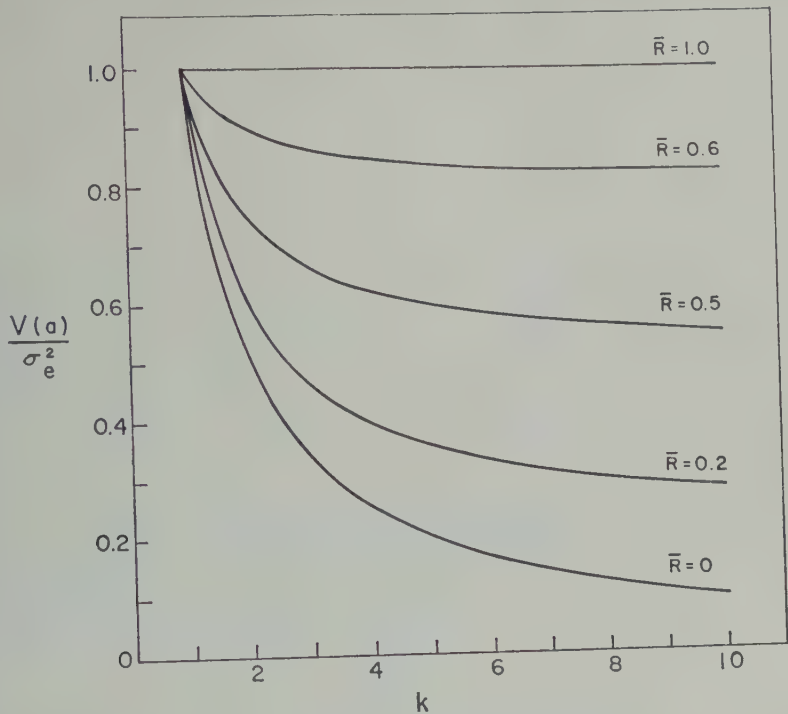


Fig. 1. Variance of a regression constant *a*.

$$= \sigma_e^2 \left\{ \frac{1}{\sum_{i=1}^k (x_i - \bar{x})^2} - \frac{2 \sum_{j=i+1}^k \sum_{i=1}^k R_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\left[\sum_{i=1}^k (x_i - \bar{x})^2 \right]^2} \right\} \tag{20}$$

a set of evenly spaced values of *x*, it can be shown that $\sum (x_i - \bar{x})^2 = -2 \sum \sum (x_i - \bar{x})(x_j - \bar{x})$, where $i \neq j$. By using this principle, assuming that $R_{ij} = \bar{R}$ for all values of *j*, (20) becomes

$$V(b) = \frac{\sigma_e^2 (1 - \bar{R})}{\sum_{i=1}^k (x_i - \bar{x})^2} \tag{21}$$

there is no interstation correlation, $\bar{R}_{ij} = 0$ for all values of $i \neq j$, so that (20) and (21) reduce to

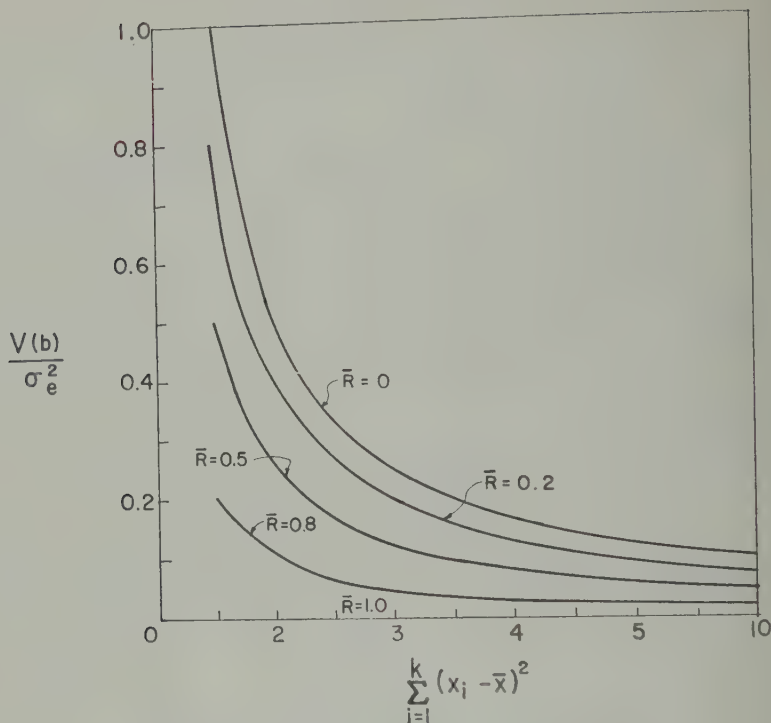
$$V(b) = \frac{\sigma_e^2}{\sum_{i=1}^k (x_i - \bar{x})^2} \tag{22}$$

Since $R_{ij} \geq 0$, interstation correlation makes the variance of *b* smaller than would be the case if there were no interstation correlation. If $R_{ij} = \bar{R} = 1$ for all values of $i \neq j$, the variance of *b* is zero. However, there may or may not be a relation between *x* and *y*. If $\bar{R} = 1$, then the set of meteorologic factors influencing the discharge of all the rivers in the given region produces either a constant increase or a constant decrease of the rivers at the same time. Thus, the relation between *x* and *y*, expressed by *b*, is not affected; however, the mean level of the classified discharge, expressed by *a*, is affected. (Tests of significance for regression coefficients when the errors are correlated have been developed by Siddiqui [1960].)

In Figure 2 a family of curves is shown for $V(b)/\sigma_e^2$ versus $\sum_{i=1}^k (x_i - \bar{x})^2$ as a function of \bar{R} . From Figure 2 it is seen that $V(b)$ tends to zero as $\sum_{i=1}^k (x_i - \bar{x})^2$ tends to infinity for all values of \bar{R} .

Regression estimates. Let *x*₀ denote a given value of *x*. If this value of *x*₀ is substituted into (2), the estimated value of *y* becomes

$$\hat{y} = a + b(x_0 - \bar{x}) \tag{23}$$

Fig. 2. Variance of the regression coefficient b .

The expectation of \hat{y} is given by

$$E(\hat{y}) = \alpha + \beta(x_0 - \bar{x}) \quad (24)$$

The variance of \hat{y} may be expressed as

$$\begin{aligned} V(\hat{y}) &= E[\hat{y} - \alpha - \beta(x_0 - \bar{x})]^2 \\ &= E[(a - \alpha) + (b - \beta)(x_0 - \bar{x})]^2 \end{aligned} \quad (25)$$

whereby

$$\begin{aligned} V(\hat{y}) &= E(a - \alpha)^2 + (x_0 - \bar{x})^2 E(b - \beta)^2 \\ &\quad + 2(x_0 - \bar{x})E[(a - \alpha)(b - \beta)] \end{aligned} \quad (26)$$

By using (6) and (16), $E[(a - \alpha)(b - \beta)]$ can be expressed as

$$\begin{aligned} E[(a - \alpha)(b - \beta)] &= E \left[\frac{1}{k} \sum_{i=1}^k e_i \frac{\sum_{i=1}^k e_i (x_i - \bar{x})}{\sum_{i=1}^k (x_i - \bar{x})^2} \right] \end{aligned} \quad (27)$$

Since e is independent of $(x - \bar{x})$, equation 27 becomes

$$\begin{aligned} E[(a - \alpha)(b - \beta)] &= \frac{\sigma_e^2}{k \sum_{i=1}^k (x_i - \bar{x})^2} \left[\sum_{j=i+1}^k \sum_{i=1}^{k-1} R_{ij}(x_i - \bar{x}) \right. \\ &\quad \left. + \sum_{j=i+1}^k \sum_{i=1}^{k-1} R_{ij}(x_j - \bar{x}) \right] \end{aligned} \quad (28)$$

If $R_{ij} = \bar{R}$ for all values of $i \neq j$, equation 28 reduces to

$$E[(a - \alpha)(b - \beta)] = 0 \quad (29)$$

By use of equation 29, equation 26 becomes

$$V(\hat{y}) = E(a - \alpha)^2 + (x_0 - \bar{x})^2 E(b - \beta)^2 \quad (30)$$

$E(a - \alpha)^2$ is the variance of a , and $E(b - \beta)^2$ is the variance of b . Thus, from (12) and (20), equation 30 becomes

$$\begin{aligned} V(\hat{y}) &= \frac{\sigma_e^2}{k} [1 + (k - 1)\bar{R}] \\ &\quad + \frac{\sigma_e^2(1 - \bar{R})(x_0 - \bar{x})^2}{\sum_{i=1}^k (x_i - \bar{x})^2} \end{aligned} \quad (31)$$

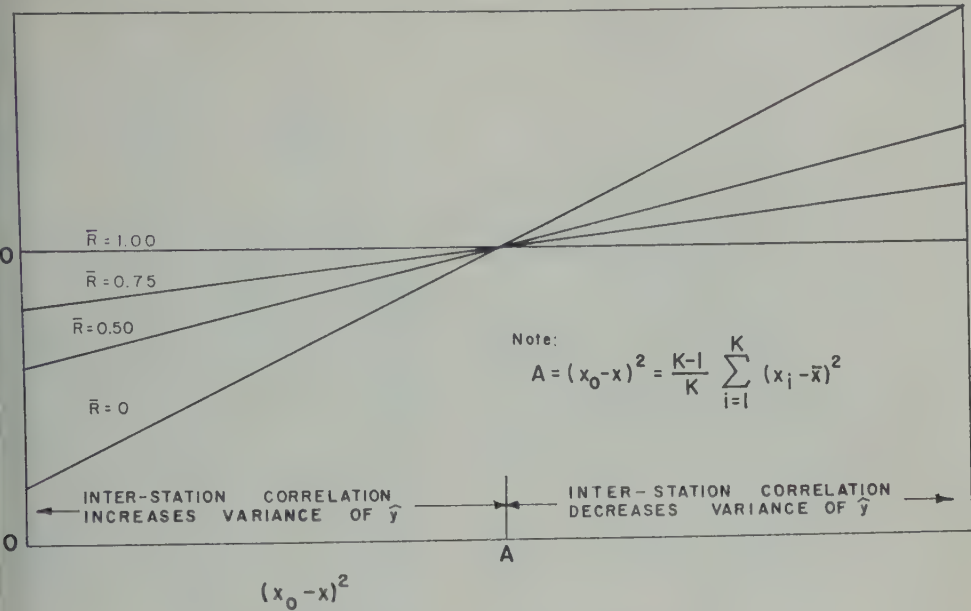


Fig. 3. Variance of a regression estimate \hat{y} .

is no interstation correlation, $\bar{R} = 0$, equation 31 then becomes

$$V(\hat{y}) = \sigma_e^2 \left[\frac{1}{k} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^k (x_i - \bar{x})^2} \right] \tag{32}$$

the variance of a increases and the effect of b decreases with an increase in \bar{R} , the effect of interstation correlation on the variance of \hat{y} is not clearly defined. By equating (32) and solving for $(x_0 - \bar{x})^2$, it is

$$(x_0 - \bar{x})^2 = \frac{(k-1)}{k} \sum_{i=1}^k (x_i - \bar{x})^2 \tag{33}$$

Equation 33 gives the value of $(x_0 - \bar{x})$ for which the variance of \hat{y} is the same when there is no interstation correlation and when there is interstation correlation. Using (33) in (32), we see that for this condition, $V(\hat{y}) = \sigma_e^2$. Moreover, $(x_0 - \bar{x})^2 > [(k-1)/k] \sum_{i=1}^k (x_i - \bar{x})^2$, the variance of \hat{y} is less when there is interstation correlation than when there is no interstation correlation. If $(x_0 - \bar{x})^2 < [(k-1)/k] \sum_{i=1}^k (x_i - \bar{x})^2$, the contrary is true. These results are shown in Figure 3, in which a family of curves for different values of \bar{R} is given for $V(\hat{y})/\sigma_e^2$ as a function of $(x_0 - \bar{x})^2$.

Example. The expressions for variance that were developed in this study have been applied to the results of a regression analysis in which the momentary peak discharges, the dependent variable, were intercorrelated in order that the effect of that intercorrelation on the results might be examined.

In a research study of regional flood-frequency relations in New England [Benson, 1961] 164 basins were used in the analysis. Five or six major storms were responsible for the three or four highest flood peaks at most of the 164 stations. Flood peaks at each station were found to be correlated to some degree with peaks at most other stations within New England. There were 13,366 possible combinations of the 164 stations. By random selection, a sample of 200 pairs was selected and was used to obtain a distribution of distances between stations and the median distance, 94 miles. A further sampling, stratified for distance, provided 54 pairs of stations. The correlation coefficients between the stations in each of the 54 pairs were computed and plotted against the distances between the stations. A coefficient of correlation of 0.26 corresponded to the median distance of 94 miles, and it was taken to be the general correlation coefficient between the 164 stations.

Alexander [1954] pointed out that where inter-

correlation exists, the equivalent number of independent stations, N' , equals

$$\frac{k}{1 + \bar{R}(k - 1)}$$

where k is the number of stations and \bar{R} is the mean correlation coefficient. For the New England data, N' equals 3.8. It is interesting to note that, assuming the same \bar{R} , if only 20 stations had been available, N' would have been 3.4, and if as many as 500 stations had been used, N' would have been 3.8. This illustrates the rapid arrival at the limiting number of independent records. No appreciable improvement in the variance of a is attained by using 500 instead of 20 stations, if they are all within the same general area.

For the New England data, a simple regression between the mean annual flood and drainage area size was found to be

$$\log Q = -.211 + .854 (\log A - \overline{\log A})$$

The error variance σ_e^2 was found to be 0.0530 (log units). The variance of a , $V(a)$, assuming noncorrelated data, would be computed by (13) as .000323. For correlated data, with \bar{R} equal to 0.26, by (12),

$$V(a) = \frac{\sigma_e^2}{k} [1 + (k - 1)\bar{R}] = \frac{\sigma_e^2}{N'}$$

$V(a)$ is computed as .014 log units, which is considerably larger than the value that would be computed for uncorrelated data.

The variance of b , from (21), is

$$V(b) = \frac{\sigma_e^2(1 - \bar{R})}{\sum_{i=1}^k (x_i - \bar{x})^2}$$

For the New England data, the denominator is equal to 62.82. For an \bar{R} of 0 (no correlation), $V(b)$ is equal to .00084, and for an \bar{R} of 0.26, $V(b)$ is equal to .00062. The variance has been reduced by the presence of the intercorrelation.

The variance of the predicted value of Q is given by (31) or

$$V(\hat{y}) = V(a) + V(b)(x_0 - \bar{x})^2$$

The variance of \hat{y} , computed for values of x_0 at 1, 2, and 3 standard deviations from the mean \bar{x} (equal to 2.176) is as follows (1 standard deviation equals 0.621):

x_0	$V(\hat{y})$	
	$\bar{R} = 0$	$\bar{R} = .26$
2.797	0.00065	0.014
3.418	.0016	.015
4.039	.0032	.016

Equation 33 gives the value of $(x_0 - \bar{x})^2$ which the variance of \hat{y} is the same with or without interstation correlation. The variance of \hat{y} (correlated) is decreased from that of (uncorrelated) when $(x_0 - \bar{x})^2$ exceeds the sum of the squares of the departures of all x 's in the original data. This is highly improbable unless x_0 is considerably beyond the range of the original set of x 's. For the New England data this would not occur unless a prediction was being made for an x_0 about 13 standard deviations from the mean. In practical applications, the variance of the predicted value is increased usually by a considerable amount, over its value with no correlation.

SUMMARY AND CONCLUSIONS

The assumptions underlying regression analyses are as follows: (1) the independent variables represent fixed values and hence do not have probability distributions; (2) the residuals from the line of regression are normally distributed; (3) the variance about the line of regression is constant; and (4) the values of the dependent variable are mutually independent. Hydrologic data, however, seldom satisfy these assumptions. By proper selection of the data, the first assumption can be met. The second and third assumptions can be satisfied by means of transformation of the dependent and/or independent variables. The fourth assumption cannot be satisfied by transformation of the data.

In regional analyses the use of regression theory is impaired by the fact that hydrologic data do not satisfy the fourth assumption. The regional problem in which this condition arises is that of relating classified values of discharge from given drainage areas within a region to physical parameters characterizing the given drainage areas. Since the rivers fluctuate more or less in response to a common set of meteorologic factors, the classified values of discharge from each of the rivers are correlated. These correlations are referred to as interstation correlations.

investigate the effect of interstation correlation on regression analysis it was assumed the regression model was linear and had one independent variable. Moreover, it was assumed that the independent variable was fixed, the residuals about the line of regression were normally distributed, that the variance about the line of regression was a constant, and the variance of each of the classified values was equal to a constant. Under these assumptions the effect of interstation correlation was studied by determining the following statistics: comparing those values with the values for the statistics when there is no interstation correlation: (1) the expectation of the regression coefficient, a ; (2) the variance of a ; (3) the expectation of the regression coefficient, b ; (4) the variance of b ; (5) the expectation of an estimated value of the dependent variable, y ; and (6) the variance of \hat{y} .

In least-squares regression analysis is made, interstation correlation does not affect the estimation of a , b , and \hat{y} . These estimates are unbiased estimators of the population values of the statistics. The variances of these statistics, however, are affected by the interstation correlation.

The variance of a is larger when there is interstation correlation than when there is no interstation correlation. If there is no interstation correlation, the variance of a goes to zero as the number of streams, k , in the region goes to infinity. However, if there is interstation correlation, the variance of a goes to $\bar{R}\sigma_e^2$ as the number of streams goes to infinity, where \bar{R} is the average of all possible interstation correlations and σ_e^2 is the variance of the residuals about the line of regression.

The variance of b is smaller when there is interstation correlation than when there is no interstation correlation. Assume that b is tested for significance and that the interstation correlation is not taken into account. If b is found to be significant, then b would be significant if the interstation correlation were taken into consideration. However, if b is found to be insignificant it is necessary to consider the interstation correlation before making an inference. Depending on the magnitude of the interstation correlation, b may or may not be significant. The variance of \hat{y} is affected by interstation correlation. However, the variance of \hat{y} may be

less or greater than would be the case if there were no interstation correlation. The value of \hat{y} pertains to a value x_0 of the independent variable. If $(x_0 - \bar{x})^2 < [(k-1)/k] \sum_{i=1}^k (x_i - \bar{x})^2$, the variance of \hat{y} is larger when there is interstation correlation than when there is no interstation correlation. If $(x_0 - \bar{x})^2 > [(k-1)/k] \sum_{i=1}^k (x_i - \bar{x})^2$, the contrary is true.

Although interstation correlation does not affect the expected values of a , b , and \hat{y} , it does affect their variance. Hence, in order that valid inferences may be made from a regression analysis, the interstation correlation must be taken into consideration when it is present.

When the effect of interstation correlation on the prediction of floods is considered, it must be concluded that we cannot predict floods within narrow confidence bands if by this we mean predicting the long-term T -year peak discharge either at a gaged or ungaged site, from the usual length of records available. On the other hand, we can use all the data within the period of record as though they represented the true distribution, and can attain consistent if not absolute standards for design purposes by differentiating the peak discharges on various basins by means of the physical and climatic characteristics.

The conclusions arrived at here are entirely general. They apply to any process in which a hydrologic variable, be it rainfall, evaporation, or any other phenomenon for which the data are interrelated on a probability or time basis, is being related to a set of pertinent variables. They also apply to much that is outside the field of hydrology.

Acknowledgment. We wish to thank Dr. J. R. Rosenblatt of the National Bureau of Standards, Washington, D. C., for reviewing the derived statistics.

Publication authorized by Director of U. S. Geological Survey.

REFERENCES

- Alexander, G. N., Some aspects of time series in hydrology, *J. Inst. Engrs. Australia*, 26, 196, 1954.
- Benson, M. A., Factors influencing the occurrence of floods in a humid region of diverse terrain: *U. S. Geol. Survey Water Supply Paper 1580-B*, in process, 1961.
- Siddiqui, M. M., Tests for regression coefficients when errors are correlated, *Ann. Math. Statistics*, 13, 929-938, 1960.

(Manuscript received April 1, 1961;
revised July 21, 1961.)

The Frequency Distribution of Near Extremes

L. E. BORGMAN

*Department of Statistics, University of California
Berkeley, California*

Abstract. The probability that the near-extreme occurrence of a physical phenomenon (e.g., the largest rainfall, the second largest flood) will exceed a selected value is important in many geophysical problems. A simplified technique for computing this probability which is applicable to a wide class of problems is presented. Utilized in the method is a tabulation of the percentage points for the exact frequency distribution of the m th largest of n values ($m = 1, 2, 3, 4, 5; m \leq n \leq \infty$) in terms of the parameter $n[1 - F(x)]$. The procedure can be used for samples either of fixed or random sizes and is particularly applicable if (1) data are missing or unreliable, (2) the sample size is small, or (3) it is desired to check an assumed population distribution against the observed near extremes. The technique is illustrated by various applications.

1. INTRODUCTION

Questions related to the extremes of natural phenomena frequently arise in geophysical investigations. The 'largest' rainfall, the 'worst' flood, and the 'most severe' hurricane are all of this type. The various techniques evolved for answering these questions range from empirical methods to statistical procedures based on various idealized assumptions. Perhaps most prominent in the latter category are the methods proposed by Gumbel [1954, 1958]. Despite the wide class of problems which can be treated satisfactorily by these techniques, situations occasionally occur for which they are unsuitable. Examples are:

(1) The maximum observations may be missing or unreliable although their relative rank can be estimated. In almost all of the methods for near extremes (e.g., highest river stage during the year) are used. However, it is precisely during these occurrences that the data are most likely to be lost or imprecise. For example, the recording devices may go off scale or the record paper may run out without replacement.

(2) The limiting distributions assumed in many current analytical techniques may not be appropriate because n may not be sufficiently large to permit the approximate use of one of the limiting forms. In such cases the exact distribution is more suitable for use in the analysis.

(3) The questions involve second largest or third largest. The common techniques give results relative only to the largest occurrence. However, the design problems motivating the investigation may require, in addition, estimates for the k th largest.

(4) The number of occurrences of the phenomenon in a fixed interval of time may itself be a random variable. The assumption of fixed sample size implicit in many of the current analysis procedures then no longer holds.

The following methods may be used to overcome these difficulties in many applications.

2. FORMULAS AND TABLES FOR SAMPLES OF FIXED SIZE

Assume that n independent observations are made in a population having a time independent frequency distribution $F(x)$ (i.e., $F(x) = \Pr\{X \leq x\}$).¹ The n observations of X are placed in order of size as $t_1 \geq t_2 \geq \dots \geq t_n$. The symbol t_m thus denotes the m th largest of the n observations (t_1 is the largest x observed, t_2 the second largest, etc.). Let $G_{m,n}(x)$ be the probability that, for the n observations, t_m is less than or equal to x . Now the m th largest observation is less than or equal to x if and only if at most $m - 1$ of the observations are greater

¹ $\Pr\{ \}$ denotes the probability of the event contained within the braces.

TABLE 1. Analysis of Example B

		From G_1 Table				From $G_2(x)$ Table		
Rank	Approx. P	$w_{100}(x)$	$F(x)$	Assoc. x		$w_{100}(x)$	$F(x)$	Assoc. x
(Largest)	1	16/17 = .941	0.061	0.99939	Missing	0.389	0.99611	32.9
	2	15/17 = .882	.126	.99874	33.3	.586	.99414	32.5
	3	14/17 = .823	.196	.99804	33.0	.758	.99242	32.4
	4	13/17 = .765	.270	.99730	Missing	.92	.9908	32.3
	5	12/17 = .706	.348	.99652	32.9	1.08	.9892	32.3
	6	11/17 = .647	.438	.99562	32.7	1.24	.9876	32.0
	7	10/17 = .588	.532	.99468	Missing	1.41	.9859	32.0
	8	9/17 = .529	.639	.99361	32.6	1.58	.9842	32.0
	9	8/17 = .471	.755	.99245	32.5	1.77	.9823	32.0
	10	7/17 = .412	.885	.99115	32.5	1.97	.9803	32.0
	11	6/17 = .353	1.05	.9895	32.4	2.20	.9780	31.9
	12	5/17 = .294	1.22	.9878	32.4	2.45	.9755	31.7
	13	4/17 = .235	1.46	.9854	32.2	2.77	.9723	31.7
	14	3/17 = .176	1.76	.9824	32.2	3.17	.9683	31.6
	15	2/17 = .118	2.16	.9784	31.9	3.67	.9633	31.5
(Smallest)	16	1/17 = .059	2.83	.9717	31.8	4.51	.9549	31.5

than x . For if m or more of the observations are greater than x , then certainly t_m is greater than x . Similarly, if less than m observations are greater than x , then the m th largest must be less than or equal to x . Each of the n observations can be considered as a binomial trial with probability $F(x)$ that it be less than or equal to x , and probability $1 - F(x)$ that it be greater than x . Hence

$$G_{m,n}(x)$$
$$= \Pr \{ \text{at most } m - 1 \text{ observations} > x \}$$
$$= \sum_{k=0}^{m-1} \binom{n}{k} [1 - F(x)]^k [F(x)]^{n-k} \tag{1}$$

The series in (1) may be expressed in terms of the incomplete β function [Pearson, 1924; Feller, 1950, pp. 162-163, (108)] as

$$G_{m,n}(x) = \binom{n}{m} m \int_0^{F(x)} t^{n-m}(1-t)^{m-1} dt \tag{2}$$

where $\binom{n}{m}$ is defined as

$$\binom{n}{m} = \frac{n(n-1)(n-2) \cdots (n-m+1)}{m!} \tag{3}$$

The above proof closely parallels that given by Schmetterer [1956, p. 345, theorem 1]. In the foregoing, no restriction was placed on $F(x)$ other than that it be a distribution function. $F(x)$ may not be differentiable, for example.

If m and $F(x)$ are held fixed in (1) as n increases indefinitely, $G_{m,n}(x)$ tends to zero. This follows from the relation $\lim_{n \rightarrow \infty} n^a \theta^n = 0$ for $0 \leq \theta < 1$ [Rudin, 1953, p. 43, theorem 3.20(d)], since with $\theta = F(x)$

$$G_{m,n}(x) = \theta^n$$
$$= \sum_{k=0}^{m-1} \frac{n(n-1) \cdots (n-k+1)}{k!} \left(\frac{1-\theta}{\theta} \right)^k$$
$$\leq \sum_{k=0}^{m-1} [\theta^n n^k] \frac{1}{k!} \left(\frac{1-\theta}{\theta} \right)^k$$
$$\rightarrow 0 \quad \text{as } n \rightarrow \infty \tag{4}$$

This is inconvenient, since it would be desirable to have a stable nonzero limiting function that could be used for large n and be listed versus $n = \infty$ in tables. A stable limit is obtained if following Cramér [1945, pp. 370-371], the parameter $w_n(x) = n[1 - F(x)]$ is used instead of $F(x)$. Let $\tilde{G}_{m,n}(w_n)$ be the expression obtained by substituting $w_n(x)$ into $G_{m,n}(x)$. That is

$$\tilde{G}_{m,n}(w_n) = \left(1 - \frac{w_n}{n} \right)^n$$
$$= \sum_{k=0}^{m-1} \binom{n}{k} \left[\frac{\frac{w_n}{n}}{1 - \frac{w_n}{n}} \right]^k = G_{m,n}(x) \tag{5}$$

holding $w_n = w$ fixed (and recalling that $\lim_{n \rightarrow \infty} (1 + 1/n)^n = e$) we obtain, as $n \rightarrow \infty$

$$\left(1 - \frac{w_n}{n}\right)^n \rightarrow e^{-w} \quad (6)$$

$$\frac{\left(\frac{w_n}{n}\right)^k}{k!} \frac{(n-1) \cdots (n-k+1)}{n^k} \rightarrow \frac{w^k}{k!} \left(1 - \frac{w}{n}\right)^k$$

$$\frac{\left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)}{1 - \frac{w}{n}} \frac{w^k}{k!} \rightarrow \frac{w^k}{k!} \quad (7)$$

as $n \rightarrow \infty$ with $w_n = w$ fixed

$$\bar{G}_{m,n}(w_n) \rightarrow e^{-w} \sum_{k=0}^{m-1} \frac{w^k}{k!} = \bar{G}_{m,\infty}(w) \quad (8)$$

is expressed in terms of the incomplete gamma function [Feller, 1950, p. 163, (10.10)] and w_n substituted into (2)

$$m \binom{n}{m} \int_0^{w_n/n} (1-r)^{n-m} r^{m-1} dr \quad (9)$$

$$1 - \frac{1}{(m-1)!} \int_0^w e^{-s} s^{m-1} ds \quad (10)$$

values of m , n , and $P = \bar{G}_{m,n}(w_n)$ are specified, the value of w_n satisfying (9) and (10) is a fixed number, although it may be somewhat difficult to compute. Tables 2 to 6 solve this problem and give w_n (in the body of the table) for $P = .005, .025, .050, .100(.1), .900, .950, .975, .990$, $m = 1, 2, 3, 4, 5$; and $m \leq n \leq \infty$. The values are given to three significant figures are taken off from more complete tables with six significant figures [Borgman, 1959].

3. FIXED SAMPLE SIZE APPLICATIONS

Example A: $F(x)$ given, $G_{m,n}(x)$ to be computed. Suppose that under the usual day-to-day regime of a certain river, the daily maximum river stage for each of the next 100 days is normally distributed with an average value of 30 feet

and a standard deviation of 2 feet. Suppose that, barring major storms, which occur only very rarely, the stage each day is independent of that for the preceding day. Levees constructed of local fill parallel the river bank. Excluding the very rare major storms, the main damage to the levees on this river is associated with erosion occurring when the river level exceeds 34 feet. The levees are breached and become ineffective if 3 days with maximum river stages exceeding 34 feet occur before maintenance and repair. We will assume that a maintenance inspection has just been made and that the next inspection is scheduled for 100 days from now. Under the usual daily regime of the river (i.e., excluding major storms), what is the probability that the levees will become ineffective before being repaired?

Solution. Here x is the daily maximum river stage. There are 100 observations of x , so $n = 100$. The levees become ineffective if and only if the third largest of these observations exceeds 34 feet. Thus

Pr {levees become ineffective}

$$\begin{aligned} &= \Pr \{t_3 > 34\} \\ &= 1 - G_{3,100}(34) \\ &= 1 - \bar{G}_{3,100}(w_{100}) \end{aligned}$$

where $w_{100} = 100 [1 - F(34)]$. Since $F(x)$ is the normal distribution with a mean of 30 feet and a standard deviation of 2 feet, it follows from the standard normal tables that $F(34) = 0.9772$. Hence $w_{100} = 100[1 - 0.9772] = 2.28$, and

$$\begin{aligned} \text{Pr \{levees become ineffective\}} \\ &= 1 - \bar{G}_{3,100}(2.28). \end{aligned}$$

Table 3 gives $\bar{G}_{3,100}(2.28) = 0.600$, so finally

$$\begin{aligned} \text{Pr \{levees become ineffective\}} \\ &= 1.0 - 0.600 = 0.400. \end{aligned}$$

Since this probability is fairly large, more frequent inspections would seem to be required. For example, maintenance once a month gives $n = 30$, $w_n = 30(.0228) = 0.684$, $1 - \bar{G}_{3,30}(0.684) = 1 - 0.969 = 0.031$, so the probability that the levees become ineffective in a 30-day period is only .031.

In the above problem, $F(x)$ was given as time independent. In an actual application, it would be necessary either to verify that this assumption is reasonable or to recognize that

TABLE 2. Percentage Points for the Frequency Distribution of the Largest of n Values
($n[1 - F(x)]$ is tabulated in the body of the table.)

n	.005	.010	.025	.050	.100	.200	.300	.400	.500	.600	.700	.800	.900	.950	.975	.990	.995
1	.995	.990	.975	.950	.900	.800	.700	.600	.500	.400	.300	.200	.100	.050	.025	.010	.005
2	1.86	1.80	1.68	1.55	1.37	1.11	.905	.735	.586	.451	.300	.200	.100	.051	.025	.010	.005
3	2.49	2.35	2.12	1.89	1.61	1.25	.992	.790	.619	.470	.336	.215	.103	.051	.025	.010	.005
4	2.94	2.74	2.41	2.11	1.75	1.33	1.04	.819	.636	.480	.341	.217	.104	.051	.025	.010	.005
5	3.27	3.01	2.61	2.25	1.85	1.38	1.07	.837	.647	.486	.344	.218	.104	.051	.025	.010	.005
6	3.52	3.22	2.76	2.36	1.91	1.41	1.09	.850	.655	.490	.346	.219	.104	.051	.025	.010	.005
7	3.72	3.37	2.87	2.44	1.96	1.44	1.11	.859	.660	.493	.348	.220	.105	.051	.025	.010	.005
8	3.87	3.50	2.96	2.50	2.00	1.46	1.12	.866	.664	.495	.349	.220	.105	.051	.025	.010	.005
9	4.00	3.60	3.03	2.55	2.03	1.47	1.13	.871	.667	.497	.350	.220	.105	.051	.025	.010	.005
10	4.11	3.69	3.08	2.59	2.06	1.49	1.13	.876	.670	.498	.350	.221	.105	.051	.025	.010	.005
12	4.28	3.82	3.18	2.65	2.10	1.51	1.15	.882	.674	.500	.351	.221	.105	.051	.025	.010	.005
14	4.41	3.92	3.24	2.70	2.12	1.52	1.15	.887	.676	.502	.352	.221	.105	.051	.025	.010	.005
16	4.51	4.00	3.29	2.73	2.14	1.53	1.16	.891	.678	.503	.353	.222	.105	.051	.025	.010	.005
18	4.59	4.06	3.34	2.76	2.16	1.54	1.16	.893	.680	.504	.353	.222	.105	.051	.025	.010	.005
20	4.65	4.11	3.37	2.78	2.17	1.55	1.17	.896	.681	.504	.354	.222	.105	.051	.025	.010	.005
25	4.77	4.21	3.43	2.82	2.20	1.56	1.18	.900	.684	.506	.354	.222	.105	.051	.025	.010	.005
30	4.86	4.27	3.47	2.85	2.22	1.57	1.18	.902	.685	.507	.355	.222	.105	.051	.025	.010	.005
40	4.96	4.35	3.52	2.89	2.24	1.58	1.19	.906	.687	.508	.355	.223	.105	.051	.025	.010	.005
50	5.03	4.40	3.56	2.91	2.25	1.58	1.19	.908	.688	.508	.355	.223	.105	.051	.025	.010	.005
60	5.07	4.43	3.58	2.92	2.26	1.59	1.19	.909	.689	.509	.356	.223	.105	.051	.025	.010	.005
80	5.13	4.48	3.61	2.94	2.27	1.59	1.19	.911	.690	.509	.356	.223	.105	.051	.025	.010	.005
100	5.16	4.50	3.62	2.95	2.28	1.60	1.20	.912	.691	.510	.356	.223	.105	.051	.025	.010	.005
140	5.20	4.53	3.64	2.96	2.28	1.60	1.20	.913	.691	.510	.356	.223	.105	.051	.025	.010	.005
180	5.22	4.55	3.65	2.97	2.29	1.60	1.20	.914	.692	.510	.356	.223	.105	.051	.025	.010	.005
220	5.24	4.56	3.66	2.98	2.29	1.60	1.20	.914	.692	.510	.356	.223	.105	.051	.025	.010	.005
260	5.24	4.56	3.66	2.98	2.29	1.60	1.20	.915	.692	.510	.356	.223	.105	.051	.025	.010	.005
300	5.25	4.57	3.67	2.98	2.29	1.61	1.20	.915	.692	.510	.356	.223	.105	.051	.025	.010	.005
350	5.26	4.58	3.67	2.98	2.30	1.61	1.20	.915	.692	.510	.356	.223	.105	.051	.025	.010	.005
400	5.26	4.58	3.67	2.98	2.30	1.61	1.20	.915	.693	.510	.357	.223	.105	.051	.025	.010	.005
500	5.27	4.58	3.68	2.99	2.30	1.61	1.20	.916	.693	.511	.357	.223	.105	.051	.025	.010	.005
700	5.28	4.59	3.68	2.99	2.30	1.61	1.20	.916	.693	.511	.357	.223	.105	.051	.025	.010	.005
1000	5.28	4.59	3.68	2.99	2.30	1.61	1.20	.916	.693	.511	.357	.223	.105	.051	.025	.010	.005
2000	5.29	4.60	3.69	2.99	2.30	1.61	1.20	.916	.693	.511	.357	.223	.105	.051	.025	.010	.005
∞	5.30	4.61	3.69	3.00	2.30	1.61	1.20	.916	.693	.511	.357	.223	.105	.051	.025	.010	.005

P

n	.005	.010	.025	.050	.100	.200	.300	.400	.500	.600	.700	.800	.900	.950	.975	.990	.995
2	1.99	1.99	1.97	1.95	1.90	1.79	1.67	1.55	1.41	1.26	1.10	.894	.632	.447	.316	.200	.141
3	2.88	2.82	2.72	2.59	2.41	2.14	1.91	1.70	1.50	1.30	1.09	.861	.587	.406	.283	.177	.124
4	3.56	3.44	3.22	3.01	2.72	2.33	2.03	1.78	1.54	1.32	1.09	.849	.570	.390	.270	.168	.118
5	4.07	3.89	3.58	3.29	2.92	2.45	2.11	1.82	1.57	1.33	1.09	.839	.561	.382	.264	.163	.114
6	4.43	4.23	3.85	3.49	3.06	2.53	2.16	1.86	1.59	1.34	1.09	.839	.556	.377	.260	.161	.112
7	4.79	4.50	4.05	3.64	3.17	2.60	2.20	1.88	1.60	1.34	1.09	.837	.552	.374	.257	.159	.111
8	5.05	4.72	4.21	3.77	3.25	2.64	2.23	1.90	1.61	1.35	1.09	.835	.549	.371	.255	.157	.110
9	5.26	4.90	4.34	3.86	3.32	2.68	2.25	1.91	1.62	1.35	1.09	.834	.547	.369	.253	.156	.109
10	5.44	5.04	4.45	3.94	3.37	2.71	2.27	1.92	1.62	1.35	1.09	.833	.545	.368	.252	.155	.109
12	5.72	5.27	4.62	4.06	3.45	2.76	2.30	1.94	1.63	1.36	1.09	.831	.543	.366	.250	.154	.108
14	5.94	5.45	4.74	4.15	3.51	2.79	2.32	1.95	1.64	1.36	1.09	.830	.541	.364	.249	.153	.107
16	6.10	5.58	4.84	4.22	3.55	2.81	2.33	1.96	1.64	1.36	1.09	.829	.540	.363	.248	.153	.107
18	6.23	5.69	4.91	4.28	3.59	2.83	2.34	1.97	1.65	1.36	1.09	.828	.539	.362	.248	.152	.106
20	6.34	5.78	4.97	4.32	3.62	2.85	2.35	1.97	1.65	1.36	1.09	.827	.538	.361	.247	.152	.106
25	6.54	5.94	5.09	4.40	3.67	2.88	2.37	1.98	1.66	1.37	1.10	.827	.537	.360	.246	.151	.105
30	6.68	6.05	5.17	4.46	3.71	2.90	2.38	1.99	1.66	1.37	1.10	.826	.536	.359	.245	.151	.105
40	6.86	6.19	5.26	4.53	3.75	2.92	2.40	2.00	1.66	1.37	1.10	.826	.535	.358	.245	.150	.105
50	6.97	6.28	5.32	4.57	3.78	2.94	2.40	2.00	1.67	1.37	1.10	.826	.534	.358	.244	.150	.104
60	7.04	6.33	5.36	4.60	3.80	2.94	2.41	2.01	1.67	1.37	1.10	.825	.533	.357	.243	.149	.104
80	7.14	6.41	5.42	4.63	3.82	2.96	2.42	2.01	1.67	1.37	1.10	.825	.533	.357	.243	.149	.104
100	7.20	6.45	5.45	4.66	3.83	2.96	2.42	2.01	1.67	1.37	1.10	.825	.533	.357	.243	.149	.104
140	7.26	6.51	5.48	4.68	3.85	2.97	2.43	2.01	1.67	1.37	1.10	.825	.533	.356	.243	.149	.104
180	7.30	6.54	5.50	4.69	3.86	2.98	2.43	2.02	1.68	1.37	1.10	.825	.532	.356	.243	.149	.104
220	7.32	6.55	5.51	4.70	3.86	2.98	2.43	2.02	1.68	1.38	1.10	.825	.532	.356	.243	.149	.104
260	7.34	6.57	5.52	4.71	3.87	2.98	2.43	2.02	1.68	1.38	1.10	.825	.532	.356	.243	.149	.104
300	7.35	6.58	5.53	4.71	3.87	2.99	2.43	2.02	1.68	1.38	1.10	.825	.532	.356	.242	.149	.104
350	7.36	6.59	5.54	4.72	3.87	2.99	2.43	2.02	1.68	1.38	1.10	.825	.532	.356	.242	.149	.104
400	7.37	6.59	5.54	4.72	3.88	2.99	2.43	2.02	1.68	1.38	1.10	.825	.532	.356	.242	.149	.104
500	7.38	6.60	5.55	4.73	3.88	2.99	2.44	2.02	1.68	1.38	1.10	.824	.532	.356	.242	.149	.104
700	7.40	6.61	5.55	4.73	3.88	2.99	2.44	2.02	1.68	1.38	1.10	.824	.532	.355	.242	.149	.104
1000	7.41	6.62	5.56	4.73	3.88	2.99	2.44	2.02	1.68	1.38	1.10	.824	.532	.355	.242	.149	.104
2000	7.42	6.63	5.57	4.74	3.89	2.99	2.44	2.02	1.68	1.38	1.10	.824	.532	.355	.242	.149	.103
∞	7.43	6.64	5.57	4.74	3.89	2.99	2.44	2.02	1.68	1.38	1.10	.824	.532	.355	.242	.149	.103

TABLE 4. Percentage Points for the Frequency Distribution of the Third Largest of n Values
($n[1 - F(x)]$ is tabulated in the body of the table.)

n	P														
	.005	.010	.025	.050	.100	.200	.300	.400	.500	.600	.700	.800	.900	.950	.995
3	2.99	2.99	2.97	2.95	2.90	2.78	2.66	2.53	2.38	2.21	2.01	1.75	1.39	1.00	.513
4	3.88	3.83	3.73	3.61	3.43	3.15	2.91	2.68	2.46	2.22	1.97	1.67	1.28	.994	.646
5	4.59	4.47	4.27	4.05	3.77	3.37	3.05	2.77	2.50	2.23	1.95	1.63	1.23	.946	.776
6	5.14	4.96	4.66	4.37	4.00	3.51	3.14	2.82	2.53	2.24	1.94	1.61	1.21	.919	.733
7	5.58	5.35	4.97	4.61	4.17	3.62	3.21	2.86	2.55	2.24	1.93	1.60	1.19	.901	.709
8	5.94	5.65	5.21	4.80	4.31	3.70	3.26	2.89	2.56	2.25	1.93	1.59	1.17	.889	.693
9	6.23	5.91	5.40	4.95	4.41	3.76	3.30	2.92	2.58	2.25	1.93	1.58	1.17	.880	.682
10	6.48	6.12	5.56	5.07	4.50	3.81	3.33	2.94	2.59	2.26	1.93	1.58	1.16	.873	.674
12	6.88	6.45	5.81	5.26	4.63	3.89	3.38	2.96	2.60	2.26	1.92	1.57	1.15	.862	.667
14	7.17	6.70	5.99	5.40	4.72	3.94	3.41	2.98	2.61	2.26	1.92	1.56	1.14	.855	.658
16	7.40	6.89	6.14	5.50	4.79	3.98	3.44	3.00	2.62	2.27	1.92	1.56	1.14	.850	.648
18	7.59	7.04	6.25	5.58	4.85	4.01	3.46	3.01	2.62	2.27	1.92	1.56	1.13	.846	.644
20	7.74	7.17	6.34	5.65	4.90	4.04	3.47	3.02	2.63	2.27	1.92	1.55	1.13	.843	.641
25	8.03	7.40	6.51	5.78	4.98	4.09	3.50	3.04	2.64	2.27	1.92	1.55	1.12	.838	.637
30	8.22	7.56	6.62	5.86	5.03	4.12	3.52	3.05	2.64	2.27	1.92	1.55	1.12	.834	.634
40	8.47	7.76	6.77	5.97	5.11	4.16	3.54	3.06	2.65	2.28	1.92	1.54	1.11	.828	.630
50	8.63	7.89	6.86	6.03	5.15	4.18	3.56	3.07	2.66	2.28	1.92	1.54	1.11	.826	.627
60	8.73	7.97	6.92	6.07	5.18	4.20	3.57	3.08	2.66	2.28	1.92	1.54	1.11	.824	.626
80	8.86	8.08	6.99	6.13	5.21	4.22	3.58	3.08	2.66	2.28	1.91	1.54	1.11	.823	.624
100	8.94	8.14	7.04	6.16	5.23	4.23	3.59	3.09	2.67	2.28	1.91	1.54	1.11	.823	.623
140	9.04	8.22	7.09	6.20	5.26	4.24	3.59	3.09	2.67	2.28	1.91	1.54	1.11	.821	.622
180	9.09	8.26	7.12	6.22	5.27	4.25	3.60	3.10	2.67	2.28	1.91	1.54	1.10	.820	.622
220	9.12	8.28	7.14	6.23	5.28	4.26	3.60	3.10	2.67	2.28	1.91	1.54	1.10	.820	.621
260	9.15	8.30	7.15	6.24	5.29	4.26	3.60	3.10	2.67	2.28	1.91	1.54	1.10	.820	.620
300	9.16	8.32	7.16	6.25	5.29	4.26	3.61	3.10	2.67	2.28	1.91	1.54	1.10	.819	.620
350	9.18	8.33	7.17	6.26	5.30	4.27	3.61	3.10	2.67	2.28	1.91	1.54	1.10	.819	.620
400	9.19	8.34	7.18	6.26	5.30	4.27	3.61	3.10	2.67	2.28	1.91	1.54	1.10	.819	.620
500	9.21	8.35	7.19	6.27	5.30	4.27	3.61	3.10	2.67	2.28	1.91	1.54	1.10	.819	.620
700	9.23	8.37	7.20	6.28	5.31	4.27	3.61	3.10	2.67	2.28	1.91	1.54	1.10	.818	.619
1000	9.24	8.38	7.21	6.28	5.31	4.27	3.61	3.10	2.67	2.28	1.91	1.54	1.10	.818	.619
2000	9.26	8.39	7.22	6.29	5.32	4.28	3.61	3.10	2.67	2.28	1.91	1.54	1.10	.818	.619
∞	9.27	8.41	7.22	6.30	5.32	4.28	3.62	3.11	2.67	2.29	1.91	1.54	1.10	.818	.619

n	P																
	.050	.010	.025	.050	.100	.200	.300	.400	.500	.600	.700	.800	.900	.950	.975	.990	.995
4	3.99	3.99	3.97	3.95	3.90	3.78	3.66	3.52	3.36	3.18	2.96	2.67	2.25	1.89	1.59	1.26	1.06
5	4.89	4.84	4.74	4.62	4.44	4.16	3.91	3.67	3.43	3.18	2.89	2.55	2.08	1.71	1.42	1.11	.925
6	5.60	5.49	5.29	5.08	4.79	4.39	4.06	3.76	3.47	3.18	2.86	2.49	2.00	1.63	1.34	1.04	.862
7	6.18	6.00	5.71	5.42	5.05	4.55	4.16	3.82	3.50	3.18	2.84	2.45	1.95	1.58	1.29	.996	.824
8	6.64	6.41	6.04	5.69	5.24	4.67	4.24	3.87	3.52	3.18	2.82	2.43	1.92	1.54	1.26	.968	.799
9	7.03	6.75	6.31	5.90	5.39	4.76	4.30	3.90	3.54	3.18	2.81	2.41	1.89	1.52	1.23	.947	.781
10	7.35	7.03	6.52	6.07	5.52	4.84	4.34	3.93	3.55	3.18	2.81	2.39	1.88	1.50	1.22	.932	.768
12	7.86	7.47	6.86	6.33	5.70	4.95	4.41	3.97	3.57	3.19	2.80	2.38	1.85	1.47	1.19	.911	.749
14	8.25	7.79	7.11	6.52	5.84	5.03	4.46	4.00	3.59	3.19	2.79	2.36	1.83	1.46	1.17	.896	.736
16	8.55	8.05	7.30	6.67	5.94	5.09	4.50	4.02	3.60	3.19	2.79	2.35	1.82	1.44	1.16	.886	.727
18	8.79	8.25	7.46	6.78	6.02	5.14	4.53	4.04	3.60	3.19	2.79	2.35	1.81	1.43	1.15	.878	.720
20	8.99	8.41	7.58	6.87	6.08	5.17	4.55	4.05	3.61	3.20	2.78	2.34	1.80	1.43	1.15	.872	.715
25	9.36	8.72	7.80	7.04	6.20	5.24	4.60	4.08	3.62	3.20	2.78	2.33	1.79	1.41	1.13	.862	.706
30	9.61	8.93	7.96	7.16	6.28	5.29	4.62	4.09	3.63	3.20	2.78	2.33	1.78	1.41	1.13	.855	.700
40	9.93	9.20	8.15	7.30	6.38	5.34	4.66	4.11	3.64	3.20	2.77	2.32	1.77	1.40	1.12	.847	.693
50	10.14	9.36	8.27	7.39	6.44	5.38	4.68	4.13	3.65	3.21	2.77	2.31	1.77	1.39	1.11	.842	.688
60	10.27	9.47	8.35	7.45	6.48	5.40	4.69	4.13	3.65	3.21	2.77	2.31	1.76	1.38	1.11	.839	.686
80	10.44	9.61	8.46	7.53	6.53	5.43	4.71	4.14	3.66	3.21	2.77	2.30	1.76	1.38	1.10	.835	.682
100	10.55	9.70	8.52	7.57	6.56	5.45	4.72	4.15	3.66	3.21	2.77	2.30	1.76	1.38	1.10	.832	.680
140	10.67	9.80	8.59	7.62	6.59	5.47	4.73	4.16	3.66	3.21	2.77	2.30	1.75	1.37	1.10	.830	.678
180	10.74	9.85	8.63	7.65	6.61	5.48	4.74	4.16	3.67	3.21	2.77	2.30	1.75	1.37	1.10	.828	.677
220	10.78	9.89	8.65	7.67	6.63	5.48	4.74	4.16	3.67	3.21	2.77	2.30	1.75	1.37	1.09	.827	.676
260	10.81	9.91	8.67	7.68	6.63	5.49	4.75	4.17	3.67	3.21	2.76	2.30	1.75	1.37	1.09	.827	.675
300	10.83	9.93	8.68	7.69	6.64	5.49	4.75	4.17	3.67	3.21	2.76	2.30	1.75	1.37	1.09	.826	.674
350	10.85	9.94	8.70	7.70	6.65	5.50	4.75	4.17	3.67	3.21	2.76	2.30	1.75	1.37	1.09	.825	.674
400	10.87	9.96	8.70	7.71	6.65	5.50	4.75	4.17	3.67	3.21	2.76	2.30	1.75	1.37	1.09	.825	.674
500	10.89	9.97	8.72	7.72	6.66	5.50	4.75	4.17	3.67	3.21	2.76	2.30	1.75	1.37	1.09	.825	.673
700	10.92	9.99	8.73	7.73	6.66	5.51	4.76	4.17	3.67	3.21	2.76	2.30	1.75	1.37	1.09	.824	.673
1000	10.93	10.01	8.74	7.74	6.67	5.51	4.76	4.17	3.67	3.21	2.76	2.30	1.75	1.37	1.09	.824	.673
2000	10.96	10.03	8.75	7.74	6.67	5.51	4.76	4.17	3.67	3.21	2.76	2.30	1.75	1.37	1.09	.824	.673
∞	10.98	10.05	8.77	7.75	6.68	5.52	4.76	4.18	3.67	3.21	2.76	2.30	1.74	1.37	1.09	.823	.672

TABLE 6. Percentage Points for the Frequency Distribution of the Fifth Largest of n Values
($n[1 - F(x)]$ is tabulated in the body of the table.)

n	P																
	.005	.010	.025	.050	.100	.200	.300	.400	.500	.600	.700	.800	.900	.950	.975	.990	.995
5	4.99	4.99	4.97	4.95	4.90	4.78	4.66	4.51	4.35	4.16	3.93	3.62	3.15	2.75	2.39	1.99	1.73
6	5.89	5.84	5.74	5.62	5.44	5.16	4.91	4.66	4.41	4.14	3.84	3.47	2.94	2.51	2.15	1.77	1.52
7	6.61	6.50	6.31	6.10	5.81	5.40	5.07	4.76	4.45	4.14	3.79	3.38	2.83	2.39	2.03	1.65	1.42
8	7.20	7.03	6.74	6.46	6.08	5.57	5.18	4.82	4.48	4.13	3.76	3.33	2.76	2.31	1.96	1.59	1.36
9	7.69	7.46	7.09	6.74	6.29	5.71	5.26	4.87	4.50	4.13	3.74	3.29	2.71	2.26	1.91	1.54	1.31
10	8.09	7.82	7.38	6.96	6.46	5.81	5.32	4.91	4.52	4.13	3.73	3.27	2.67	2.22	1.87	1.50	1.28
12	8.73	8.37	7.81	7.31	6.71	5.96	5.42	4.96	4.54	4.13	3.71	3.23	2.62	2.17	1.82	1.46	1.24
14	9.21	8.78	8.13	7.56	6.89	6.07	5.49	5.00	4.56	4.13	3.69	3.21	2.59	2.14	1.79	1.43	1.21
16	9.59	9.10	8.38	7.75	7.02	6.15	5.54	5.03	4.57	4.13	3.68	3.19	2.57	2.11	1.76	1.41	1.19
18	9.89	9.36	8.57	7.90	7.13	6.22	5.58	5.06	4.58	4.13	3.68	3.18	2.55	2.10	1.75	1.39	1.18
20	10.13	9.57	8.73	8.02	7.21	6.27	5.61	5.07	4.59	4.14	3.67	3.17	2.54	2.08	1.73	1.38	1.17
25	10.59	9.95	9.02	8.24	7.37	6.36	5.67	5.11	4.61	4.14	3.66	3.15	2.52	2.06	1.71	1.36	1.15
30	10.90	10.21	9.22	8.39	7.47	6.42	5.70	5.13	4.62	4.14	3.66	3.14	2.50	2.04	1.69	1.34	1.13
40	11.30	10.54	9.47	8.58	7.60	6.49	5.75	5.16	4.63	4.14	3.65	3.13	2.48	2.02	1.67	1.33	1.12
50	11.55	10.75	9.62	8.69	7.68	6.54	5.78	5.17	4.64	4.14	3.65	3.12	2.47	2.01	1.66	1.32	1.11
60	11.72	10.89	9.72	8.77	7.73	6.57	5.80	5.18	4.64	4.14	3.65	3.11	2.47	2.00	1.66	1.31	1.11
80	11.93	11.06	9.85	8.86	7.80	6.61	5.82	5.20	4.65	4.14	3.64	3.11	2.46	2.00	1.65	1.30	1.10
100	12.06	11.17	9.93	8.92	7.83	6.63	5.83	5.20	4.66	4.14	3.64	3.10	2.45	1.99	1.64	1.30	1.09
140	12.21	11.29	10.02	8.99	7.88	6.66	5.85	5.21	4.66	4.15	3.64	3.10	2.45	1.98	1.64	1.29	1.09
180	12.30	11.36	10.07	9.02	7.91	6.67	5.86	5.22	4.66	4.15	3.64	3.10	2.44	1.98	1.63	1.29	1.09
220	12.35	11.41	10.10	9.05	7.92	6.68	5.87	5.22	4.66	4.15	3.64	3.10	2.44	1.98	1.63	1.29	1.09
260	12.39	11.44	10.12	9.06	7.93	6.69	5.87	5.22	4.66	4.15	3.64	3.09	2.44	1.98	1.63	1.29	1.08
300	12.41	11.46	10.14	9.08	7.94	6.69	5.87	5.23	4.67	4.15	3.64	3.09	2.44	1.98	1.63	1.28	1.08
350	12.44	11.48	10.15	9.09	7.95	6.69	5.87	5.23	4.67	4.15	3.64	3.09	2.44	1.98	1.63	1.28	1.08
400	12.46	11.49	10.16	9.09	7.95	6.70	5.88	5.23	4.67	4.15	3.64	3.09	2.44	1.98	1.63	1.28	1.08
500	12.49	11.52	10.18	9.11	7.96	6.70	5.88	5.23	4.67	4.15	3.63	3.09	2.44	1.97	1.63	1.28	1.08
700	12.52	11.54	10.20	9.12	7.97	6.71	5.88	5.23	4.67	4.15	3.63	3.09	2.44	1.97	1.63	1.28	1.08
1000	12.54	11.56	10.21	9.13	7.98	6.71	5.88	5.23	4.67	4.15	3.63	3.09	2.43	1.97	1.63	1.28	1.08
2000	12.57	11.58	10.23	9.14	7.99	6.72	5.89	5.24	4.67	4.15	3.63	3.09	2.43	1.97	1.62	1.28	1.08
∞	12.59	11.60	10.24	9.15	7.99	6.72	5.89	5.24	4.67	4.15	3.63	3.09	2.43	1.97	1.62	1.28	1.08

ulas and tables give only simplified
ations to the actual situation.

le B: Data given, $F(x)$ to be computed.

that we desire to check the normality
ons in example A and have as data the
argest and many of the largest observa-
the last 16 of the 100-day intervals. The
data, with x indicating missing (but
ately ranked) information is as follows.
t observations: x , 33.3, 33.0, x , 32.9,
32.6, 32.5, 32.5, 32.4, 32.4, 32.2, 32.2,
8.

l largest observations: 32.9, 32.5, 32.4,
3, 32.0, 32.0, 32.0, 32.0, 31.9, 31.7,
6, 31.5, 31.5.

m. As a quick check, we note that
55) ≈ 0.5 (i.e., the median of the largest
ions is 32.55). But the $w_{100}(x)$ correspond-
 $G_{1,100}(x) = 0.5$ is $w_{100}(x) = 0.691$ or
 $1 - 0.691/100 = 0.993$. If x were
with mean 30 and standard deviation 2,
for $F(x) = 0.993$ would be given by
 $b)/2 = 2.45$ or $x = 34.9$. This is quite
from $x = 32.55$, which was observed
suggests a significant deviation from the
y assumption. The most direct way to
to find the correct $F(x)$ is by reconstruc-
follows. With each observed largest and
argest, we associate the empirical esti-
 $G_{1,100}(x)$ and $G_{2,100}(x)$ given by $1 -$
1). Here r is the rank counted from the
to the smallest, and N is the number of
ions (16 in the example, Table 1). The
his particular estimate is suggested by
[1954, pp. 13-15]. The value of $w_{100}(x)$
ed with each of the empirical $G_1(x)$ and
ay then be obtained from the tables and
s shown in Table 1. Finally, $F(x)$ is
ed from $F(x) = 1 - w_{100}(x)/100$. A
f $F(x)$ versus x , both as derived from the
and from the second-largest observations,
rovides an estimate of the important
ail of the distribution. Table 1 gives the
cal results for the specified example, and
1 shows the graph of $F(x)$ as a function
e see that $F(x)$ is approximately normal
ean 29.2 feet but that the standard
on is approximately 1.4 instead of 2 feet.
 x), thus computed, would appear to be
tly accurate for problems of the upper
es but would not be very reliable outside
er tail.

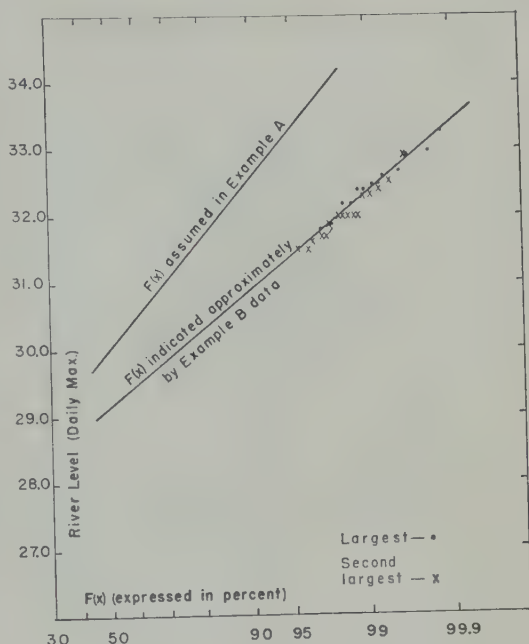


Fig. 1. Results for example B.

4. FORMULAS FOR SAMPLES OF RANDOM SIZE

In many problems, the number of occurrences
of the phenomenon in a unit of time is itself
random. (For example, the number of Atlantic
hurricanes in the coming year is random.) The
previously developed formulas are no longer
applicable. To accommodate this new situation,
we shall construct a statistical model which,
though somewhat elaborate, will be found to
have the desirable properties of generality and
versatility.

Let Δt denote the basic time interval in the
application. (If the problem involves the third
largest hurricane in a 10-year period, then
 $\Delta t = 10$ yr.) Each occurrence of the phenomenon
in a selected future Δt period will be called an
event and the (random) number of events in
 Δt will be indicated by N . The lower-case letter
 n will be used to specify a particular realization
of the random variable N . This same convention
of capital letters for random variables and
lower-case letters for realizations will also be
used for the other random quantities of the
model (excluding t_m which was defined at the
start of section 2).

Suppose that out of all the N events that occur
in Δt , we are interested only in those that have

some specified property (e.g., out of all Gulf of Mexico hurricanes, we are interested only in those that pass within 75 miles of Galveston, Texas). The set of all N events will be called the totality of events, and the S events having the specified property will be called the successful events. The number S is also a random variable, and for any realizations of N and S , $0 \leq s \leq n$. It will be assumed that each of the totality of events can be considered as an independent binomial trial, with probability ϕ that it is a successful event. Thus

$$\begin{aligned} \Pr \{S = s \text{ given } N = n\} \\ = \binom{n}{s} \phi^s (1 - \phi)^{n-s} \end{aligned} \quad (11)$$

Each of the totality of events is assumed to have an intensity random variable X associated with it (e.g., for a hurricane, this might be the maximum ground level wind speed). The distribution function of X will be denoted by $F(x) = \Pr \{X \leq x\}$.

In most applications of extremal statistics, one is concerned with the occurrence or nonoccurrence of some type of disaster. The disaster is considered to have occurred if, for specified values of m and x , the intensity t_m of the m th largest of the successful events exceeds x . If less than m successful events occur, there cannot be an m th event and no disaster occurs. Therefore it is natural to consider only the following mutually exclusive and exhaustive possibilities:

$$A. \quad S \geq m \quad \text{and} \quad t_m > x \quad (\text{disaster})$$

$$B. \quad S < m, \quad \text{or both} \quad S \geq m$$

$$\text{and } t_m \leq x \quad (\text{no disaster})$$

The symbol $D_m(x)$ will be used to indicate the probability that the disaster (i.e., A) occurs. Hence

$$D_m(x) = \Pr \{S \geq m \quad \text{and} \quad t_m > x\} \quad (12)$$

The fundamental problem for the model is the determination of a formula for $D_m(x)$.

Let $\Pr \{N = n\}$ be denoted by $P_N(n)$. From (11)

$$\begin{aligned} \Pr \{N = n \quad \text{and} \quad S = s\} \\ = \Pr \{N = n\} \Pr \{S = s \text{ given } N = n\} \\ = P_N(n) \binom{n}{s} \phi^s (1 - \phi)^{n-s} \end{aligned} \quad (13)$$

Summing (13) over n gives $P_S(s)$ (defined $\Pr \{S = s\}$) as

$$P_S(s) = \sum_{n=s}^{\infty} P_N(n) \binom{n}{s} \phi^s (1 - \phi)^{n-s} \quad (14)$$

From (2)

$$\begin{aligned} \Pr \{t_m > x \text{ given } S = s\} &= 1 - G_{m,s}(x) \\ &= \int_F^1 \binom{s}{m} m t^{s-m} (1-t)^{m-1} dt \end{aligned} \quad (15)$$

The probability that $t_m > x$ is obtained summing $\Pr \{S = s \text{ and } t_m > x\}$ over s . Hence by (14) and (15)

$$\begin{aligned} D_m(x) &= \Pr \{t_m > x\} \\ &= \sum_{s=m}^{\infty} P_S(s) \Pr \{t_m > x \text{ given } S = s\} \\ &= \sum_{s=m}^{\infty} \sum_{n=s}^{\infty} P_N(n) \binom{n}{s} \phi^s (1 - \phi)^{n-s} \\ &\quad \cdot \int_F^1 \binom{s}{m} m t^{s-m} (1-t)^{m-1} dt \end{aligned} \quad (16)$$

The summation signs may be moved inside the integral and the order of summation interchanged:

$$\left(\sum_{s=m}^{\infty} \sum_{n=s}^{\infty} = \sum_{n=m}^{\infty} \sum_{s=m}^n \right).$$

It is also convenient to change the variable of integration to $v = \alpha(1-t)$, with α as an arbitrary constant which can later be assigned a desired value. With these changes and the definition $v_0 = \alpha[1 - F(x)]$, (16) becomes

$$\begin{aligned} D_m(x) &= \int_{v_0}^{v_0} \frac{m v^{m-1} \phi^m}{\alpha^m} \sum_{n=m}^{\infty} P_N(n) \\ &\quad \cdot \left[\sum_{s=m}^n \binom{n}{s} \binom{s}{m} (1 - \phi)^{n-s} \left(\phi - \frac{\phi v}{\alpha} \right)^{s-m} \right] dv \end{aligned} \quad (17)$$

The quantity in brackets can be considerably simplified with the binomial formula [Burlington 1949, p. 44]:

$$(a + b)^K = \sum_{j=0}^K \binom{K}{j} a^j b^{K-j} \quad (18)$$

Since

$$\binom{n}{s} \binom{s}{m} = \binom{n}{m} \binom{n-m}{n-s},$$

acketed quantity can be rewritten after

$$j = n - s \text{ as}$$

$$\begin{aligned} &= \binom{n}{m} \sum_{i=0}^{n-m} \binom{n-m}{j} (1-\phi)^i \\ &\quad \cdot \left(\phi - \frac{\phi v}{\alpha}\right)^{n-m-i} \\ &= \binom{n}{m} \left[(1-\phi) + \left(\phi - \frac{\phi v}{\alpha}\right) \right]^{n-m} \\ &= \binom{n}{m} \left(1 - \frac{\phi v}{\alpha}\right)^{n-m} \end{aligned} \quad (19)$$

ubstitution of (19) into (17) with

$$\binom{n}{m} = \frac{n!}{m! (n-m)!}$$

$$\begin{aligned} &= \int_0^{v_0} \frac{v^{m-1} \phi^m}{(m-1)! \alpha^m} \left[\sum_{n=m}^{\infty} \frac{n!}{(n-m)!} \right. \\ &\quad \cdot P_N(n) \left(1 - \frac{\phi v}{\alpha}\right)^{n-m} \Big] dv \end{aligned} \quad (20)$$

of the summations in (16) has thus been hatched. It would seem desirable to eliminate ther also, especially since it has an infinite set. This can be done if we introduce the ability-generating function $\gamma_N(t)$ for the om variable N . This function is defined to [eller, 1950, pp. 248-253] the expected value for any $0 \leq t \leq 1$. That is

$$\gamma_N(t) = E(t^N) = \sum_{n=0}^{\infty} P_N(n) t^n \quad (21)$$

he following examples show, $\gamma_N(t)$ often a fairly simple form:

0) The Poisson distribution

$$P_N(n) = e^{-\lambda} \lambda^n / n! \quad (22)$$

he λ is a constant equal to the average value Then, since

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

ington, 1949, p. 44],

$$\begin{aligned} \gamma_N(t) &= \sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n t^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \\ &= e^{-\lambda} e^{\lambda t} \\ &= e^{-\lambda(1-t)} \end{aligned} \quad (23)$$

(b) The binomial distribution

$$P_N(n) = \binom{K}{n} p^n (1-p)^{K-n} \quad (24)$$

where K is the number of trials, n is the number of 'successes,' and p is the probability of a 'success' on a single trial. Hence

$$\begin{aligned} \gamma_N(t) &= \sum_{n=0}^K \binom{K}{n} p^n (1-p)^{K-n} t^n \\ &= \sum_{n=0}^K \binom{K}{n} (pt)^n (1-p)^{K-n} \end{aligned}$$

and so by (18)

$$\gamma_N(t) = (1-p+pt)^K \quad (25)$$

(c) The negative binomial distribution

$$P_N(n) = \binom{k+n-1}{n} \theta^n / (1+\theta)^{k+n} \quad (26)$$

In this formula, k and θ are constants which can be estimated from data [Thom, 1957, pp. 8-10]. Thus

$$\gamma_N(t) = \sum_{n=0}^{\infty} \binom{k+n-1}{n} \frac{\theta^n t^n}{(1+\theta)^{k+n}}$$

This is the binomial series with negative exponent [Burlington, 1949, p. 44]; it equals

$$\gamma_N(t) = (1+\theta-\theta t)^{-k} \quad (27)$$

The quantity within the brackets in (20) is closely related to the probability-generating function of N . The m th derivative of (21) with respect to t is

$$\gamma_N^{(m)}(t) = \sum_{n=m}^{\infty} \frac{n!}{(n-m)!} P_N(n) t^{n-m} \quad (28)$$

If t in (28) is replaced by $(1-\phi v/\alpha)$, (28) becomes the bracketed quantity of (20). Thus

$$\begin{aligned} D_m(x) &= \int_0^{v_0} \frac{v^{m-1} \phi^m}{(m-1)! \alpha^m} \gamma_N^{(m)} \left(1 - \frac{\phi v}{\alpha}\right) dv \end{aligned} \quad (29)$$

The infinite summation in (20) has now been absorbed into $\gamma_N(t)$ and eliminated.

The substitution of particular $\gamma_N(t)$ into (29) often produces very elementary formulas. For example, suppose N is a Poisson random variable.

Then, from (23),

$$\begin{aligned}\gamma_N^{(m)}(t) &= \lambda^m e^{-\lambda(1-t)} \\ \gamma_N^{(m)}\left(1 - \frac{\phi v}{\alpha}\right) &= \lambda^m e^{-\lambda\phi v/\alpha}\end{aligned}\quad (30)$$

Hence from (29)

$$D_m(x) = \int_0^{v_0} \frac{v^{m-1} \phi^m}{(m-1)! \alpha^m} \lambda^m e^{-\lambda\phi v/\alpha} dv$$

Since α may be assigned any constant value, let $\alpha = \lambda\phi$. Then

$$D_m(x) = \int_0^{\lambda\phi[1-F(x)]} \frac{v^{m-1} e^{-v}}{(m-1)!} dv \quad (31)$$

This is essentially (10) with $w = \lambda\phi[1 - F(x)]$. So $1 - D_m(x)$ is given in the tables as the P value corresponding to the above-defined w value (w being in the body of the table on the line $n = \infty$).

In summary, the foregoing model consists of the following:

1. For the time interval Δt , the probability that the totality of events consists of n occurrences is $P_N(n)$. The probability-generating function of N is $\gamma_N(t)$. For a given event, the intensity X has the distribution function $F(x)$ and is independent of the intensities of the other events.

2. With each member of the totality of events there is associated an independent binomial trial which is such that there is a probability ϕ that the member is also a successful event.

3. The probability that the m th largest of the successful events exists and exceeds x is given by (29) where $v_0 = \alpha[1 - F(x)]$, where $\gamma_N^{(m)}(1 - \phi v/\alpha)$ denotes the m th derivative of $\gamma_N(t)$ evaluated at $t = 1 - \phi v/\alpha$, and where α is a free constant that may be assigned any convenient value.

The case in which every member of the totality of events is also a success can easily be obtained by setting $\phi = 1$. Formula (29) then gives the probability that the m th largest of the totality of events exceeds x .

5. RANDOM SAMPLE SIZE APPLICATIONS

Example C. Suppose that the number of days per year during which hail falls to the ground at a given location is a Poisson random variable with $\lambda = 2.36$ [Thom, 1957]. Suppose further that for every occurring hailstorm there

is a 0.2 probability that it will occur in month of September, a critical period for damage at the location. Let the average hail diameter be distributed normally (truncated) $x < 0$ with parameters $\mu = 0.2$ inch and $\sigma = 0.1$ inch. Suppose that the average hail diameter in any selected storm is independent of the average diameter of the others and that, for the purpose of the analysis, $F(x)$ may be assumed to be time independent. If the crops are damaged appreciably whenever the average hail diameter exceeds 0.25 inch twice in September, what is the probability of appreciable damage in a given year?

Solution. (1) The totality of events consists of all hail days during an interval $\Delta t = 1$ yr. The random variable N is Poisson, so $P_N(n)$ and $\gamma_N(t)$ are given by (22) and (23), with $\lambda = 2.36$. The intensity X is the average hailstone size and it has the truncated normal distribution

$$\begin{aligned}F(x) &= \frac{1}{\sqrt{2\pi}\sigma} \\ &\cdot \int_0^x \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx \Big/ \frac{1}{\sqrt{2\pi}\sigma} \\ &\cdot \int_0^\infty \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx\end{aligned}\quad (32)$$

where $\mu = 0.2$, $\sigma = 0.1$.

(2) The class of successful events consists of the days in September during which hail falls. For each hail day, there is a probability $\phi = 0.2$ that it will occur in September.

(3) The crops are damaged appreciably if the second largest intensity (i.e., average hail diameter) of the September hail days exceeds 0.25 inch. Hence $m = 2$ and

$\Pr\{\text{appreciable damage}\} = D_2(0.25)$ (33)
The formula for the Poisson case was obtained in (31), which for $m = 2$ yields

$$D_2(x) = \int_0^{\lambda\phi[1-F(x)]} v e^{-v} dv \quad (34)$$

Upon comparing (34) with (10), we see that $1 - D_2(0.25) = \bar{G}_{2,\infty}(w)$, with $w = \lambda\phi[1 - F(0.25)]$. From (32) and tables of the normal integral $F(0.25) = 0.6687/0.9772 = 0.6843$. With the specified values of λ and ϕ ,

$w = (2.36)(0.2)[1 - 0.6843] = 0.149$.
Table 2 gives $P = \bar{G}_{2,\infty}(0.149) = 0.990$; the 0.149 is entered in the body of the table on the

$= \infty$ and the 0.990 is interpolated on the peled P . Hence (33) and (35) lead to appreciable damage} $= 1 - 0.990 = 0.010$

5. SUMMARY

procedures outlined can be used to make ability statements about the first five es: The initial information may consist an assumed $F(x)$ formula, (2) data leading empirical $F(x)$ curve, or (3) past records near extremes. In any case, the procedure s of (1) the estimation of $F(x)$ if it is not y known, and (2) the use of the tables β_n to make the probability assertions. The ures outlined are quite versatile, and no ot has been made to show all possible ith the examples.

with all statistical models, it is particularly ant that the natural phenomenon involved application satisfy, at least approximately, dependence assumptions inherent in the as. In the preceding derivations there three basic independence assumptions: he intensity X of an event was assumed statistically independent of the intensities other occurring events. (2) The dison function, $F(x)$, of X was taken to be ndent of time. (3) For the model based on n sample sizes, the class of successes was ed to be obtained from the totality of y independent binomial trials, with a ability ϕ that a given member of the totality ents was also a success. There are problems ch the geophysicist must knowingly ignore rtures of the natural phenomenon from assumptions in order to simplify the on to a point where computable answers e obtained. Considerable caution is required such circumstances, and, if possible, ults of the formulas should be compared ast records of the phenomenon to deter- the significance of the departures.

le tables are actually a special form of the mplete β and γ functions. However the ular format used in the presentation is convenient for analysis of near extremes he usual tables of β and γ functions. The of the normal distribution as the parent ation in many of the examples was purely ter of convenience. Any other distribution

could have been used, since the procedures are independent of the form of $F(x)$.

Acknowledgments. I wish to thank Dr. J. Neyman and Dr. D. H. Wright for their advice and guidance. I am particularly grateful to Dr. Neyman for his suggestion that the probability-generating function be incorporated into the analysis of random sample size.

The study was prepared with the partial support of the National Science Foundation, grant G-8211. The percentage-point tables and the fixed-sample-size analysis are taken from the author's thesis prepared in partial fulfillment of the requirements for a Master of Science degree at the University of Houston, Houston, Texas. The applications and the analysis of random sample size are based on later work.

REFERENCES

- Borgman, L. E., The frequency distribution for the m th largest of n values, M.S. thesis, University of Houston, Houston, Texas, 172 pp., 1959.
- Burington, R. S., *Handbook of Mathematical Tables and Formulas*, 3rd ed., Handbook Publishers, Inc., Sandusky, Ohio, 296 pp., 1949.
- Cramér, Herald, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N. J., 575 pp., 1945.
- Feller, W., *An Introduction to Probability Theory and Its Applications*, 2nd ed., J. Wiley & Sons, New York, 461 pp., 1950.
- Gumbel, E. J., *Statistical Theory of Extreme Values and Some Practical Applications*, NBS Applied Math. Series 33, Washington, D. C., 51 pp., 1954.
- Gumbel, E. J., *Statistics of Extremes*, Columbia University Press, New York, 375 pp., 1958.
- Pearson, Karl, Note on the relationship of the incomplete beta function to the sum of the first p terms of the binomial $(a+b)^n$, *Biometrika*, 16, 202-203, 1924.
- Pearson, Karl, *Tables of the Incomplete Beta-Function*, Cambridge University Press, Cambridge, England, 494 pp., 1934.
- Rudin, Walter, *Principles of Mathematical Analysis*, McGraw-Hill Book Co., New York, 227 pp., 1953.
- Schmetterer, L., *Einführung in die Mathematische Statistik*, Springer-Verlag, Wien, Germany, 405 pp., 1956.
- Thom, H. C. S., The frequency of hail occurrence, *Tech. Rept. 3*, Advisory Committee on Weather Control, Washington, D. C., 19 pp., June, 1957.

(Manuscript received February 24, 1961; revised August 1, 1961. Presented at the Pacific Southwest Regional Meeting, American Geophysical Union, January 26, 1961, Berkeley, California.)

Reduction of Transpiration

W. J. ROBERTS

State Water Survey Division, Urbana, Illinois

Abstract. The monomolecular film technique for reducing evaporation from water surfaces has been applied to the problem of transpiration from plants. Hybrid corn grown in soil ended by various amounts of fatty alcohols such as hexadecanol has required up to 40 per cent water during its growth than control plants. It is theorized that some of the transpiration of corn plants can safely be reduced by the blocking action of molecules of hexadecanol passed through plants and deposited at the stomatal water-vapor interface. Although the blocking action could be taking place throughout the plants, testing with C^{14} hexadecanol has produced radioautographs showing activity not only in the roots and stalks but also throughout the leaves of treated corn plants. Comparative testing has been done in the greenhouse; however, one treated corn crop, grown naturally, yielded over 100 bushels per acre. Chemical tests of this corn showed no apparent effect of large doses of hexadecanol added to the roots of the plants.

The application of monomolecular films to water surfaces has fascinated scientists for the past several decades. From the original laboratory work of Langmuir [1917] to the application technique developed by Vines [1960], there has been rapid progress. Of all the chemicals tested, monolayers of two fatty alcohols, octadecanol and hexadecanol, have been found to have superior film-forming properties. Hydrologists have previously confined their efforts to creating artificial barriers at the vapor-water interface on water surfaces. In nature, there is an analogous vapor-water interface across which water is transported in vast quantities. It is located at the minute stomatal openings in the epidermis of plants. The transpiration process is not a continuous, but it is much more rapid during the day than at night. Water evaporates from the stomatal cell walls and tends to saturate the air in the intercellular space. If the air on the outside of the open stomata is not saturated with water vapor, then water will diffuse out of the intercellular space, permitting more transpiration from the stomata. A significant amount of transpiration takes place only if the stomata are open, and the epidermis of the plant leaves is covered with a layer of cutin almost impervious to water. Thus the rate of transpiration can be controlled by the condition of the stomata, whether they are open or closed, and to the number of stomatal

openings on a leaf surface. If it is assumed that the plant roots have an adequate source of water, a likely way to reduce the plant's demand for water would be to restrict the stomatal openings.

Experience with monomolecular films produced on water surfaces has suggested the theory that such a film can be generated in the stomata of plants. The film will have the ability to permit passage of oxygen and carbon dioxide across its surface but will inhibit water molecules from escaping. The research reported in this paper is of an exploratory nature and was conducted mostly on a laboratory scale.

To test the feasibility of controlling transpiration by introducing a mixture of octadecanol and hexadecanol into plant systems, the State Water Survey initiated a preliminary program in May 1959. Dr. L. C. Bliss, University of Illinois Botany Department, cooperated by providing facilities for this study. First, a study was made of corn seeds planted in various concentrations of distilled water and hexadecanol in order to determine whether germination would take place. All the seeds germinated readily in these solutions. Next, 24 clay pots, 8 inches in diameter, were filled with sterilized earth. To each of 12 pots, 0.5 gram of a mixture of powdered octadecanol and hexadecanol was added. Corn seeds were planted in all pots. The plants were allowed to grow for 4 weeks in the labora-

tory, and at the end of that time the stalks were cut directly above the surface of the soil, and their water demands were immediately gaged by potometer, a simple botanical tool which measures the water demand of individual plants. Significant data were collected from four of the control plants and from five of those grown in the soil enriched with octa-hexadecanol. Table 1 shows these data. The average values show that the corn grown in treated soil required only 62.5 per cent of the water demanded by the control plants.

The transpiration reduction effect was later tested in several ways. To determine the weight of water lost to transpiration, corn plants were grown in small plastic cups filled only with vermiculite. Groups of 8 plants each were allowed to reach the 3-leaf stage in a period of 10 days. Eight control plants were grown in vermiculite and water. Eight other plants were grown similarly except that 0.5 gram of powdered octa-hexadecanol was added to each pot. When the plants had reached the 3-leaf stage they were watered and the pots were covered with aluminum foil through which the stalks projected. Starting at 8 A.M. the following morning, the plants were weighed at intervals until 5 P.M. The stems were then cut at the

TABLE 1. Comparison of Water Demand of Treated and Untreated Plants

Controls		Treated Plants	
Sample No.	Water/Leaf Area, g/cm ²	Sample No.	Water/Leaf Area, g/cm ²
1	0.690	2	0.42
3	.403	4	.27
5	.590	6	*
7	.453	8	.27
9	*	10	.31
11	*	12	.38

* Data unreliable and therefore omitted from the table.

surface of the aluminum foil and weighed. When the plant weights were correlated with the corresponding water losses it was found that the control plants had lost an average of 10.98 grams of water per gram of plant material whereas the plants grown in vermiculite and hexadecanol had lost 6.43 grams per gram of plant material. Figure 1 shows a graph of the weight versus time. During an 18-month period 12 groups of pots containing corn were tested in this manner. The greatest savings were recorded during the summer season when greenhouse temperature

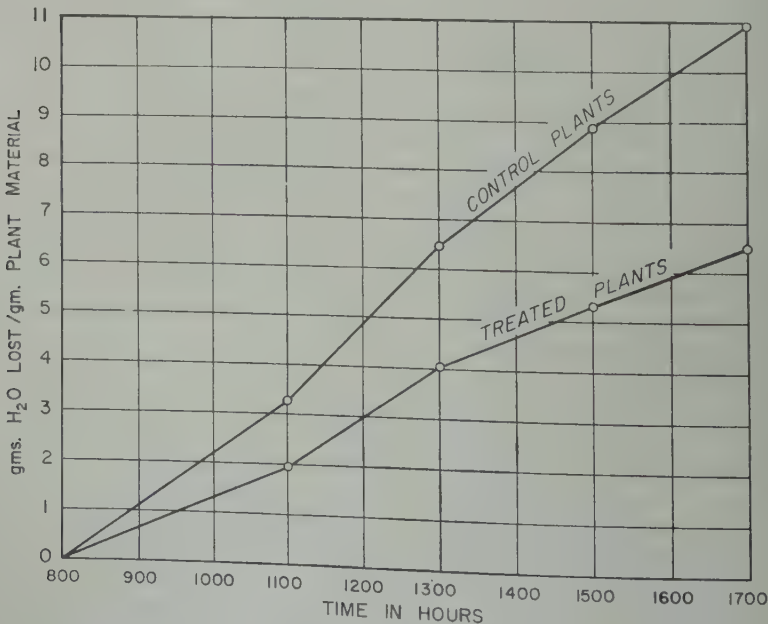


Fig. 1. Graph showing comparison of water loss from treated and control plants.



Radioautograph of corn plant grown in nutrient solution containing C^{14} hexadecanol.

were above 100°F and transpiration rates high. The lowest savings, that is those giving the least difference in water demand between treated and untreated corn grown in

pots, occurred during tests made in the winter.

To trace the route of the powdered hexadecanol through the plants a group of corn plants was grown in a nutrient solution containing radioactive hexadecanol. The plants were cut, dried, and pressed against unexposed X-ray plates for 36 hours. The resulting radioautographs indicated that the radioactive isotope had traveled through the root structure to all parts of the plant, including the leaves (see Fig. 2).

Since the radioautographs indicated a great concentration of hexadecanol in the root structures, an experiment was set up to determine the blocking effect of this chemical in the root area. Six corn plants were grown in individual pots. Two grams of powdered hexadecanol were added to the soil of three of them. All six plants were grown until the main stalks were $\frac{1}{2}$ inch in diameter. The stalks were cut and immediately subjected to nearly 1 atmosphere of vacuum for 24 hours. During that time the roots of the control plants had passed between 12 and 22 per cent more water than those grown with the hexadecanol. Thus part of the blocking action of the film-forming chemical appears to take place in the root structure (see Fig. 3).

Limited field testing was done during the 1960 growing season. A 4-acre corn field on the University of Illinois farm was made available for



Fig. 3. Vacuum installation for testing moisture movement through corn roots.

a transpiration project. Corn was planted during the first week of June 1960. On July 5, when the corn was about knee high, 1 pound of flaked hexadecanol was worked into the root area of each of 16 corn hills, which formed a square within a 2-acre area that was completely covered with a black plastic film. The corn stalks protruded through this film. The crop was harvested October 11, and a yield of more than 125 bushels per acre was obtained from the whole area. The yield from the 16 hills was equal to that of the rest of the field. After the corn was harvested an examination of the soil showed that most of the 16 pounds of hexadecanol appeared to remain in its original condition in the ground. This evidence indicates that the chemical had not affected the crop adversely and that biologic attrition of hexadecanol in soil may not be a significant problem. During the 1961 growing season the University of Illinois Agronomy Department cooperated in an extensive testing program involving 37 plots. Soil moisture tubes were installed in each plot so that accurate moisture data could be ob-

tained by the neutron scatter method. The data were not available when this paper was submitted for publication.

Water waste in nature is tremendous. Several estimates indicate that over 20 million acre-feet of water is wasted annually by nonbeneficial plants in the 17 western states. Transpiration reduction from such useless growths might release millions of acre-feet of water for beneficial purposes in these and other semiarid areas of the world.

Extrapolations based on this research provide impetus for extensive testing of the technique throughout the whole range of agriculture.

REFERENCES

- Langmuir, I., The constitution and fundamental properties of solids and liquids, 2, *Liquids*, *Am. Chem. Soc.*, 39, 1848-1906, 1917.
Vines, R. G., Evaporation control: A method of treating large storages, Manuscript read at American Chemical Society meeting, New York City, September 15, 1960.

(Manuscript received May 25, 1961;
revised August 3, 1961.)

Average Antecedent Temperatures as a Factor in Predicting Runoff from Storm Rainfall

CHARLES D. HOPKINS, JR., AND DALE O. HACKETT

*River Forecast Center, U. S. Weather Bureau
Hartford, Connecticut*

Abstract. Rainfall-runoff relations in New England and New York have been shown to vary widely from basin to basin in a manner related to average basin latitudes and elevations. Station variations and latitudes have been shown to be related to average temperatures. Average monthly and annual temperatures were computed for each basin sampled and average weekly basin temperatures estimated. These weekly temperatures were used to derive an index of average antecedent basin temperature based on a logarithmic recession. Two rainfall-runoff relationships were derived (one for the index of antecedent basin precipitation, the index of average antecedent basin temperatures corresponding to the season of the storm, the average annual basin temperature, storm rainfall, and storm runoff were used. One of these relations applied to the spring and summer, the other to the fall and winter. Testing showed that a large part of the variation in the rainfall-runoff relationship had been removed. It was concluded that average basin temperatures can be used with profit in computing runoff in New England and New York.

Statement of problem. The Hartford River Forecast Center has developed rainfall-runoff relations for basins in New England and eastern New York. Hopkins [1960] developed a standard rainfall-runoff relation for these basins in the manner described by Linsley, Kohler, and Paulhus [1949]. Examination of the data had shown that there were significant variations in the results from basin to basin. The ratio of the total predicted runoff for a given basin to the total observed runoff for the same storms, was termed 'forecast error'. This quantity is tabulated by basins in Table 1.

One objective of these endeavors was to develop rainfall-runoff relations generally applicable throughout the area of responsibility of the Hartford River Forecast Center. The 63 basins for which data were derived comprised about 25 per cent of this area. The incompleteness of these data was, therefore, a serious limitation.

It was possible, but impractical, alternative methods have been to develop a unique rainfall-runoff relationship for every forecast basin. There were, however, compelling reasons why this was not done. First, a large part of the region was not sampled. In many areas rainfall data are not sufficient to compute basin rainfall; in others the runoff data are not satisfactory. Derivation of satisfactory

runoff data can be made impossible by regulation by mills and power plants in headwater basins; in river reaches, even minor inaccuracies in ratings or routing can seriously affect the computation of runoff from the intermediate drainage.

Hopkins [1960] has shown, however, that the variations in the rainfall-runoff relations are related to mean basin elevation and mean basin latitude. Furthermore, he has demonstrated a relation between latitude, elevations, and average temperatures and has provided a tool for computing average basin temperatures.

The problem to be solved in this study was to find a way of using average basin temperatures in a rainfall-runoff relation. To define the relation it was necessary to use data from all, or at least several, basins at once. Such a relation would add a new dimension to the usual rainfall-runoff relation in that it could be said to define a way in which the relation varies from one basin to another. This would not only reconcile the data from the various sample basins but would also provide a tool with which to compute runoff from the parts of the forecast area not included in the sample.

The selection of a temperature function. It was decided to use mean temperatures in the new relation, since it had been shown that they had

TABLE 1. Comparison of Forecast Biases

Stream	Station	Number of Storms	Average Runoff	Forecast Bias,	New Runoff Rel	Fore
				Standard Relation	Average Error	
Pleasant River	Milo (nr.), Me.	41	0.457	0.45	0.145	0.1
Sheepscot River	North Whitefield, Me.	50	0.261	1.19	0.126	1.0
Dead River	The Forks, Me.	22	0.370	0.71	0.191	1.0
Diamond River	Wentworth Location (nr.), N.H.	43	0.502	0.45	0.116	0.1
Swift River	Roxbury (nr.), Me.	63	0.379	0.44	0.120	0.1
Nezinscot River	Turner Center, Me.	48	0.474	0.76	0.134	0.1
Saco River	Conway (nr.), N.H.	34	0.506	0.65	0.134	1.0
Pemigewasset River	Woodstock, N.H.	58	0.521	0.47	0.142	0.1
Baker River	Rumney (nr.), N.H.	71	0.336	0.74	0.089	1.0
Warner River	Davisville, N.H.	50	0.253	1.07	0.084	1.0
Suncook River	North Chichester, N.H.	41	0.239	1.17	0.067	1.0
So. Br. Piscataquog River	Goffstown, N.H.	56	0.283	1.11	0.075	0.9
Souhegan River	Merrimack, N.H.	55	0.341	1.13	0.084	0.1
No. Nashua River	Leominster (nr.), Mass.	61	0.315	1.24	0.083	0.9
Assabet River	Maynard, Mass.	27	0.300	1.63	0.104	1.2
Charles River	Charles River Village, Mass.	46	0.277	1.39	0.072	0.8
Neponset River	Norwood, Mass.	35	0.387	1.93	0.149	1.1
Blackstone River	Northbridge, Mass.	66	0.320	1.47	0.120	1.0
Willimantic River	South Coventry, Conn.	69	0.418	1.26	0.112	0.9
Natchaug River	Willimantic, Conn.	79	0.332	1.13	0.093	0.7
Yantic River	Yantic, Conn.	82	0.404	1.25	0.129	0.8
Upper Ammonoosuc River	Groveton (nr.), N.H.	40	0.392	0.58	0.134	1.1
Passumpsic River	Passumpsic, Vt.	61	0.285	0.70	0.080	1.1
Ammonoosuc River	Bath (nr.), N.H.	56	0.302	0.61	0.097	1.1
White River	West Hartford, Vt.	51	0.321	0.78	0.108	1.0
Mascoma River	West Canaan, N.H.	61	0.257	1.02	0.072	1.1
Sugar River	West Claremont, N.H.	43	0.244	1.02	0.064	1.1
West River	Newfane, Vt.	49	0.445	0.66	0.153	0.7
Otter Brook	Keene (nr.), N.H.	51	0.348	0.83	0.117	0.9
North River	Shuttucksville, Mass.	78	0.349	1.01	0.115	0.9
Ware River	Coldbrook, Mass.	53	0.342	1.21	0.082	0.9
W.Br. Westfield River	Huntingdon, Mass.	94	0.485	0.91	0.143	0.8
W.Br. Farmington River	Riverton (nr.), Conn.	68	0.349	1.12	0.092	1.0
Park River	Hartford, Conn.	133	0.308	1.24	0.091	0.7
Hockanum River	East Hartford, Conn.*	16	0.609	1.64	0.231	1.0
Housatonic River	Great Barrington (nr.), Mass.	51	0.384	1.03	0.092	0.9
Tenmile River	Gaylordsville (nr.), Conn.	75	0.240	1.88	0.111	1.3
Shepaug River	Roxbury (nr.), Conn.	84	0.390	1.46	0.112	1.1
Naugatuck River	Naugatuck (nr.), Conn.	86	0.376	1.28	0.095	0.9
Saugatuck River	Westport (nr.), Conn.	98	0.426	1.48	0.166	0.9
Hudson River	Cooley, N.Y.	57	0.366	0.59	0.142	1.0
Schroon River	Riverbank, N.Y.	19	0.379	1.10	0.218	1.3
Sacandaga River	Hope (nr.), N.Y.	73	0.447	0.67	0.132	0.9
Batten Kill	Battenville, N.Y.	82	0.191	1.15	0.084	1.1
Kayaderoseras Creek	West Milton (nr.), N.Y.	105	0.220	1.12	0.064	0.9
Hoosic River	Eagle Bridge (nr.), N.Y.	85	0.259	1.11	0.108	1.1
East Canada Creek	Dolgeville, N.Y.	43	0.455	0.61	0.129	0.7
East Canada Creek	East Creek, N.Y.	17	0.682	0.70	0.174	0.7
Otsquago Creek	Fort Plain, N.Y.	42	0.305	0.86	0.158	0.7
Schoharie Creek	Prattsville, N.Y.	79	0.418	0.79	0.127	0.9
Poesten Kill	Troy (nr.), N.Y.	72	0.335	0.78	0.156	0.7
Kinderhook Creek	Rossmann, N.Y.	61	0.299	1.02	0.095	0.8

TABLE 1. Continued

Stream	Station	Number of Storms	Average Runoff	Forecast	New Runoff	Relation
				Bias, Standard Relation	Average Error	Forecast Bias
Ill Creek	Oak Hill, N.Y.	79	0.304	0.89	0.144	0.92
ut Creek	Rosendale, N.Y.	71	0.395	0.93	0.175	0.74
ut Creek	Rosendale, N.Y.*	22	0.675	0.88	0.250	0.67
Ill River	Pellets Island Mt., N.Y.	77	0.380	1.01	0.155	0.67
nger Creek	Wappingers Falls (nr.), N.Y.	78	0.222	1.48	0.071	0.99
Ill Creek	Beacon, N.Y.	61	0.288	1.33	0.074	0.89
Creek	Center Rutland, Vt.	45	0.342	0.85	0.116	1.09
ski River	Montpelier, Vt.	70	0.279	0.90	0.087	1.23
iver	Northfield Falls, Vt.	69	0.357	0.70	0.132	1.00
ski River	Essex Junction (nr.), Vt.†	30	0.600	0.53	0.197	0.76
lle River	Johnson, Vt.	56	0.408	0.61	0.095	0.92
	Total	3738				
	Average		0.351	99.5	0.119	97.4
	Standard Deviation			34.4		16.2

minus area above water supply reservoir.
local area.

ter correlation with elevation and latitude
either maximum or minimum temperatures.
one of the items of storm data was the
number of the beginning date of the storm,
s necessary to have average basin tem-
ures for each week.

stants for computing average monthly
temperatures are shown in Table 2. The
icients for elevation, which had shown some
ularity, were smoothed slightly from the
s found by Hopkins [1960].

he average monthly temperatures were then
ed for each basin in terms of week number
etermining the week number for the mid-
of each month. A curve of average tem-
ure versus week number was sketched, and
these sketches of the annual march of
age basin temperature, average temperatures
estimated for each calendar week for each
l.

maximum mean basin temperatures occurred
alendar week 30, and it was already known
minimum runoff conditions occurred about
adar week 36. It was necessary, therefore,
elect a measure of antecedent temperatures
would be at a maximum at calendar week 36.
h index of antecedent temperature (*ATI*)
computed as follows:

ATI (this week) = $0.9 ATI$ (previous
week) + $0.1 \bar{T}$ (previous week)

For each basin an *ATI* was assumed for any
given week, and the computation was made
until the error involved in estimating the first
ATI had disappeared. An example of the *ATI*
and the plotting of average monthly temperatures
and the sketching of average weekly temperatures
is shown in Figure 1. It will be noted that the
ATI will not be much affected by minor in-
accuracies in estimating weekly temperatures.

TABLE 2. Constants for Computing Average
Monthly Basin Temperatures

Month	<i>a</i>	$b_{Lat},$ °F/°Lat.	$b_{Elev},$ °F/1000 ft
January	164.20	-3.25	-2.70
February	149.39	-2.90	-3.05
March	139.92	-2.45	-3.50
April	128.47	-1.90	-3.60
May	125.64	-1.55	-3.60
June	123.86	-1.30	-3.60
July	122.07	-1.15	-3.60
August	119.95	-1.15	-3.50
September	122.77	-1.40	-3.10
October	122.72	-1.65	-2.80
November	125.83	-2.00	-2.70
December	145.61	-2.75	-2.65
Annual	132.34	-1.95	-3.20
$T = a$			
+ b_{Lat} (Basin Latitude)			
+ b_{Elev} (Basin Elevation)			

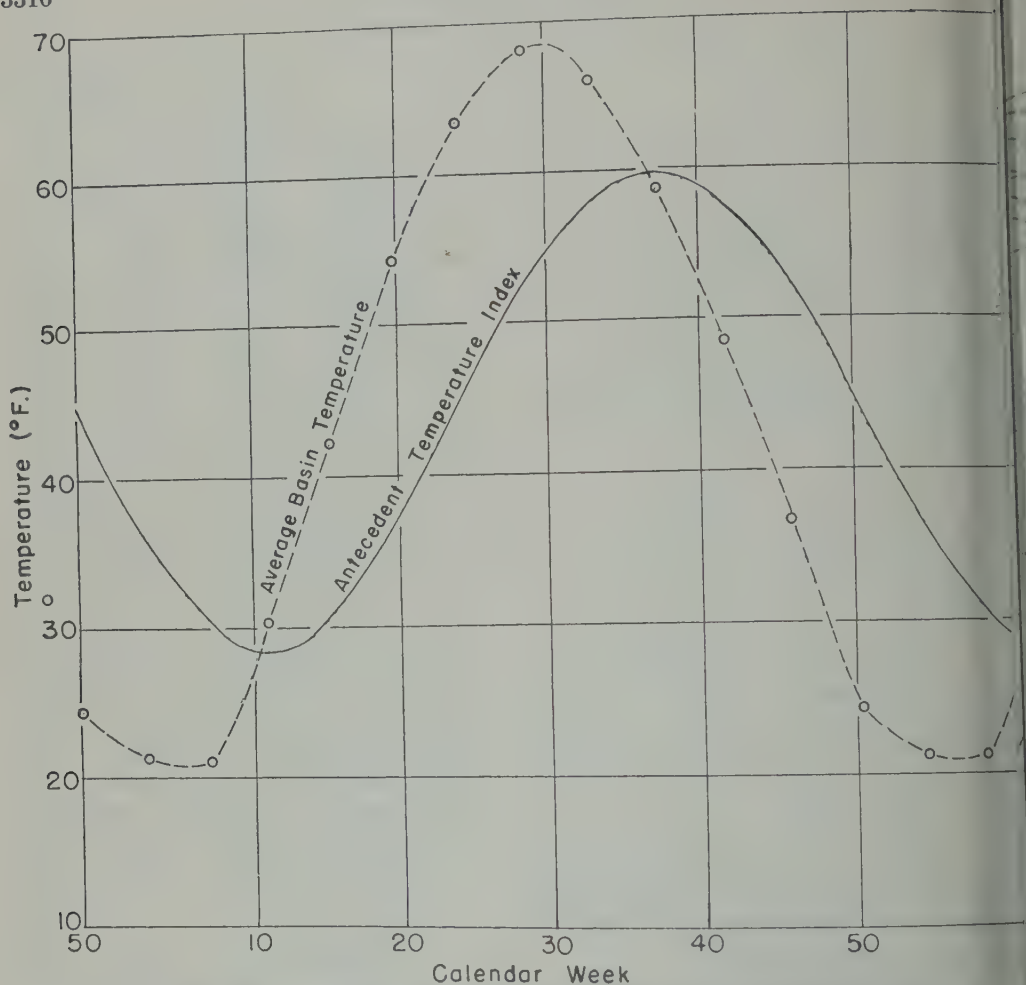


Fig. 1. Annual march of average basin temperature and average antecedent temperature index, Hoosic River basin above Eagle Bridge, N. Y.

The sample of rainfall-runoff data. The total rainfall-runoff data available consisted of 3738 storms in 63 basins. This sample was unwieldy in size because of the large number of storms; it had been intended to define the rainfall-runoff relation in each basin as well as possible.

It was necessary, therefore, to select a substantially smaller sample of workable size in order to develop a single rainfall-runoff relation. This subsample was selected according to the following criteria:

1. Every fifteenth storm.
2. Storms with precipitation greater than 3.50 inches in basins north of the latitude of the northern Massachusetts border.

3. Storms with precipitation greater than 4. inches in basins south of the latitude of the northern Massachusetts border.

The selected sample consisted of 334 storms and included storms from all the original basins. As can be seen, this selected sample included a high percentage of storms with large amounts of precipitation, and it was of a workable size. At the same time, the number of degrees of freedom afforded by this selected sample was large enough to make the results statistically significant.

Development of a rainfall-runoff relation. A rainfall-runoff relation was developed using the following variables: antecedent precipitation,

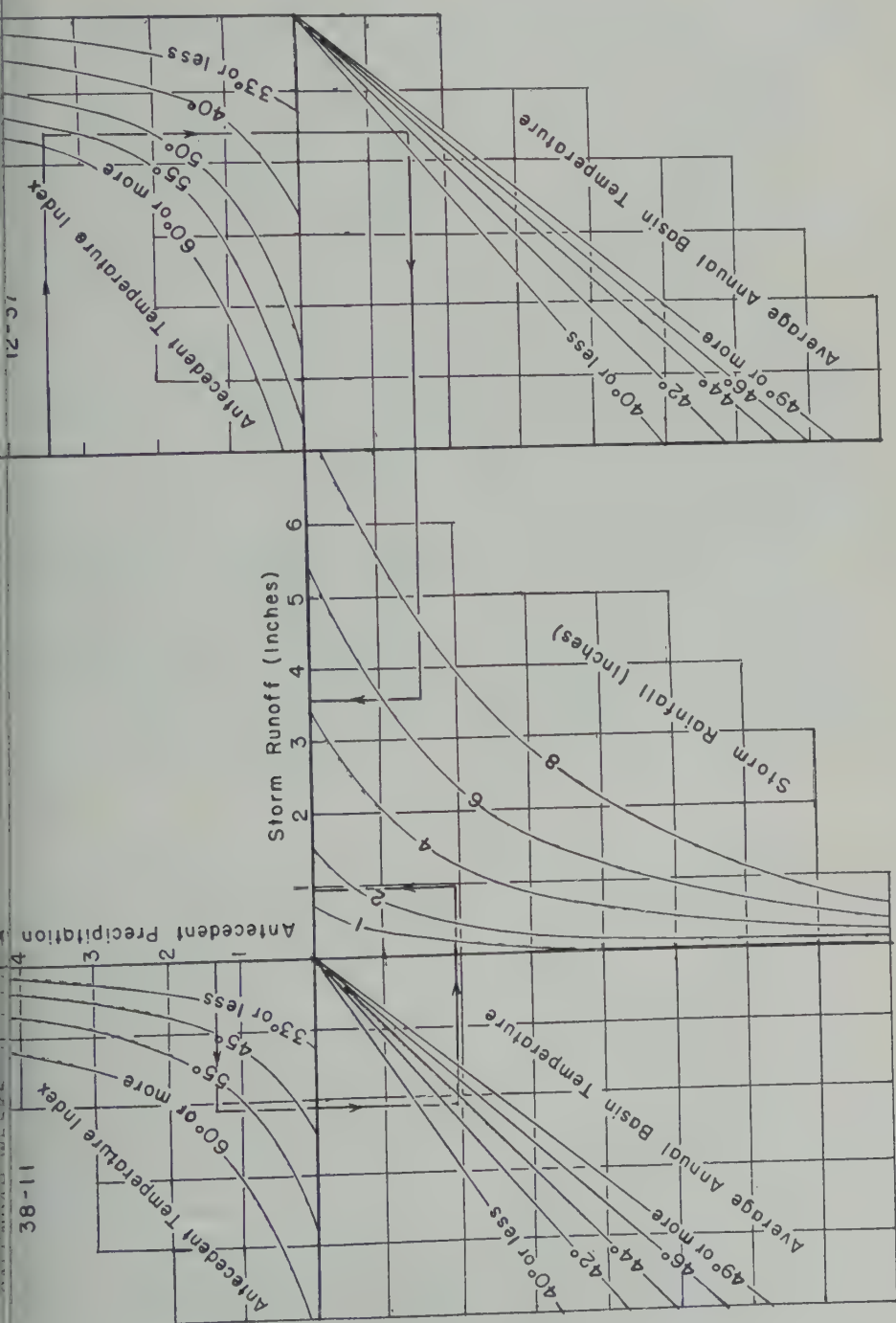


Fig. 2. A rainfall-runoff relation for New England and eastern New York based upon average basin temperatures.

index, antecedent temperature index, average annual basin temperature, storm rainfall, and storm runoff. This relation is shown in Figure 2.

The relation developed is really two separate relationships, one for the calendar weeks from calendar week 12 through week 37 and the other for calendar week 38 through week 11. The sample of 334 storms seems to be large enough to justify this differentiation.

It will be noted that, for any one basin, there is no real difference between the relation shown in Figure 2 and the standard rainfall-runoff relation shown by Linsley, Kohler, and Paulhus. This is true because, for a given basin, the antecedent temperature index is a function of calendar week. The difference between the two types of relations is, then, that in this new relation the use of average basin temperatures introduces interbasin variables. These variables provide an objective way of varying the first part of the rainfall-runoff relation according to the climates of the various basins.

Conclusions. The rainfall-runoff relation shown in Figure 2 has an average error in computing storm runoff of 0.114 inch for the total sample of 3738 storms. The standard deviation of the forecast biases between the basins, using the new relation, is 16.2 per cent. This is less than half of the standard deviation of the forecast bias using a standard rainfall-runoff relation

that was approximately the average relation for the same basins.

It can be fairly said that this new relation using average basin temperatures, explains a large part of the variations in the rainfall-runoff relations in New England and New York. Furthermore, since such average basin temperatures can be computed for any basin in the region, it provides an objective tool for estimating the rainfall-runoff relation for any basin in the region.

This relation will be used to compute runoff for all parts of the area, whether rainfall-runoff data are available or not. It is planned to adjust the computations according to the results indicated by observed storm data, not only for the basins studied, but also for neighboring or similar areas according to the dictates of judgment.

REFERENCES

- Hopkins, Charles D., Jr., A method of estimating basin temperatures in New England and New York, *J. Geophys. Research*, 65(11), 3767-3777, 1960.
- Linsley, Ray K., Jr., Max A. Kohler, and Joseph L. H. Paulhus, *Applied Hydrology*, McGraw-Hill Book Co., pp. 419-424, 1949.

(Manuscript received June 22, 1961.)

An Improved Statistical Model for Evaluating Parameters Affecting Water Yields of River Basins

B. HARRIS

Department of Mathematics, University of Nebraska, Lincoln

A. L. SHARP

U. S. Department of Agriculture, Lincoln, Nebraska

A. E. GIBBS

U. S. Department of the Interior, Lincoln, Nebraska

W. J. OWEN

U. S. Department of Agriculture, Lincoln, Nebraska

Abstract. The inference drawn from the ordinary multiple regression approach may be questionable when this method is used to analyze hydrologic data. A statistical model that avoids some of these uncertainties is developed. A Taylor series expansion is suggested to obtain exponential and interaction terms. Orthogonal transformations are used to extract some of the variables for use as predictors. A rule is exhibited for selecting the single most important (based on ability to explain variance in the dependent variable) independent variable, testing its significance, and removing its effects on all remaining variables. A second rule is exhibited for selecting, in turn, the succeeding most important independent variables, testing their significance, and removing their effects from other variables. Finally, a rule is exhibited for stopping the selection of independent variables when those remaining will not contribute significantly to the further reduction of unexplained variance in the dependent variable. The net result is the selection of a few from many independent variables to use in a 'near best' prediction equation. A method is presented for obtaining the multiple regression equation using only the selected variables. The application of the model is illustrated by its use in analyzing data from 763 storms on watershed 3H at the Central Great Plains Experimental Watershed, Hastings, Nebraska.

INTRODUCTION

The Cooperative Water Yield Procedures are a joint undertaking of the Agricultural Research Service and the Soil Conservation Service of the U. S. Department of Agriculture and the Bureau of Reclamation of the U. S. Department of Interior, the search has been continued for a statistical model more applicable than the multiple regression approach for evaluating the many parameters affecting the water yielded as stream flow by river basins. The application of the multiple regression method to this problem was explored by Sharp, Gibbs, Owen, and Harris [1960], who concluded that assumptions inherent in the method are contradicted by hydrologic data, that common tests of significance generally applied to the results derived from such methods are questionable, and that confidence may be falsely placed in estimates obtained from estimating equations

resulting from regression analyses. It was further indicated that the applicability of some of the more modern statistical techniques to the problems confronting the project should be investigated.

Basically, there are two methods available for evaluating the effects of the conservation use and treatment of land (watershed parameters) on water yielded by stream flow: (1) A comparison of treated and untreated replicated watersheds in the same climatic zone. (2) An evaluation of changes in the relation between precipitation and stream flow which occurs within a watershed as the conservation use and treatment of land progresses.

No adequate data are available from paired treated and untreated watersheds (other than from very small research ones); hence the first method enumerated above is not generally applicable. Most efforts must therefore be de-

voted to the second approach. A search for the most applicable statistical model has been conducted during all the studies performed in the Cooperative Water Yield Procedures Study project. It is believed that the model discussed later in this report offers promise of being an improved one.

THE MULTIPLE REGRESSION MODEL

In the application of multiple regression analysis to hydrologic problems, several difficulties are ordinarily encountered. These are in the selection of a model, in the choice of independent variables, in the assignment of statistical significance to the regression coefficients, and in interpreting the roles played by the independent variables as predictors of the dependent variable. Frequently, the 'shotgun' approach is attempted. All possible variables, X_1, X_2, \dots, X_p , which may be relevant and for which data are readily available, are listed, and then a linear model of the form

$$X_{0j} = \alpha + \beta_1 X_{1j} + \beta_2 X_{2j} + \dots + \beta_p X_{pj} + \epsilon_j \quad j = 1, 2, \dots, N \quad (1)$$

is assumed, where the quantities $\alpha, \beta_1, \beta_2, \dots, \beta_p$ are unknown parameters to be estimated by least squares, the ϵ_j are independent, normally distributed, random variables with mean zero and common variance σ^2 , and $X_{1j}, X_{2j}, \dots, X_{pj}$ are known constants.¹ This last assumption is intrinsically not very seriously violated, for if X_1 is the variable—average watershed precipitation—then it is understood that X_1 is not the *actual* average watershed precipitation, but is, instead, the average *measured* watershed precipitation. The difficulties are primarily in the assumption of the model and the choice of variables. For example, it may happen that $\log X_1$ and X_2 ² should be employed and that all remaining variables make no further appreciable contribution, but this combination may never be employed.

Sometimes, in order to get a regression equation involving no more than q independent variables, all possible regression equations are computed for all possible subsets of q of the p independent variables, and the equation which

minimizes the residual sum of squares is selected as a prediction equation. This approach has two principal disadvantages: first, the computation may be very time-consuming if $\binom{p}{q}$

is quite large; second, the large number of regression equations computed makes the assignment of statistical significance difficult and makes spurious results likely. The following will illustrate this. Suppose $\beta_1 = \beta_2 = \dots = \beta_p = 0$. Then if tests of the sum of squares due to regression for statistical significance at the 0.01 level are made, and $p = 20, q = 4$, we have $\binom{p}{q} = 4845$

and about 48 out of the 4845 regression equations computed would appear to be statistically significant, although not actually so.

Finally, there is the problem of assignment of statistical significance to the regression coefficients b_1, b_2, \dots, b_p . We illustrate this by a simple example when $p = 2$. The regression equation estimated by least squares is

$$X_0 = a + b_1 X_1 + b_2 X_2 \quad (2)$$

and if we were to compute simple regression equations for just X_1 or X_2 , we would have

$$X_0 = a' + b_1' X_1$$

and

$$X_0 = a'' + b_2' X_2$$

Both b_1' and b_2' may be significant, and b_1 and b_2 may not be significant. The experimenter may then be inclined to conclude (wrongly), if he computes only (2), that X_1 and X_2 do not play significant roles in predicting X_0 . This will happen if X_1 and X_2 are correlated, and if both are introduced into a regression equation their ability to explain the variability in X_0 is 'split' between them. This can be overcome by looking at (2) as a unit and not considering b_1 and b_2 separately.

THE ORTHOGONALIZATION MODEL

A method by which the data can be used to choose a model and select the variables to be used, and in which the assignment of statistical significance to the variables has a meaningful interpretation, is discussed hereafter. Further use of this method to obtain a good prediction equation using as few of the independent vari-

¹ Nomenclature used in these discussions is contained in the Appendix.

are warranted by the data is discussed. Included is a discussion of the limitations of the method, as well as its advantages, and application to predicting surface runoff from 763 storms on Experimental Watershed Hastings, Nebraska.

Methodology. Assume that there is a constant function $f(X_1, X_2, \dots, X_p)$ which has derivatives of all orders, so that

$$f(X_{1j}, X_{2j}, \dots, X_{pj}) + \epsilon_j$$

$$j = 1, 2, \dots, N \quad (3)$$

the ϵ_j are independent, normally distributed, random variables with mean zero and variance σ^2 . The specific values of the independent variables are to be chosen, and are regarded as known. Suppose, however, the exact functional relationship is unknown. $f(X_1, X_2, \dots, X_p)$ can be expanded in a series about $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p$, where

$$\bar{X}_i = \sum_j \frac{X_{ij}}{N} \quad (4)$$

N is the number of observations. This gives

$$f(X_2, \dots, X_p) = f(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p)$$

$$\sum_{i=1}^p (X_i - \bar{X}_i)$$

$$\left[\frac{\partial}{\partial X_i} f(X_1, X_2, \dots, X_p) \right]_{(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p)}$$

$$\sum_{i=1}^p \sum_{k=1}^p (X_i - \bar{X}_i)(X_k - \bar{X}_k)$$

$$\left[\frac{\partial^2}{\partial X_i \partial X_k} f(X_1, X_2, \dots, X_p) \right]_{(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p)}$$

$$\dots + R_n(X_1, X_2, \dots, X_p) \quad (5)$$

R_n is the remainder after the Taylor expansion has been carried out through partial derivatives of order n . Also, by taking n sufficiently large, R_n may be made arbitrarily small. Hence $f(X_1, X_2, \dots, X_p)$ can be approximated to within any desired degree of accuracy by a polynomial of sufficiently high degree. Let $X_{ij} = X_{ij} - \bar{X}_i$ and note that the partial derivatives are constants so that (3) may be

written

$$X_{0j} = \alpha + \sum_{i=1}^p \beta_i x_{ij} + \sum_{i=1}^p \sum_{k=1}^p \beta_{ik} x_{ij} x_{kj}$$

$$+ \dots + R_n'(x_{1j}, x_{2j}, \dots, x_{pj}) + \epsilon_j \quad (6)$$

Henceforth it will be assumed that n is so large that R_n' may be ignored, in which case (6) reduces to ordinary multiple regression with $\beta_i, \beta_{ik}, \beta_{ikl}, \dots$ as unknown parameters and with independent variables $x_i, x_{ik}, x_{ikl}, \dots$. In all there are

$$k = 1 + \binom{p}{1} + \binom{p+1}{2} + \dots$$

$$+ \binom{p+n-1}{n} = \binom{p+n}{p} \quad (7)$$

unknown parameters and $k_n - 1$ independent variables. So, to a large extent, the problem of 'guessing' the model has been eliminated, but the number of variables introduced may be very large. Accordingly, it is now desired to find a method of choosing some of the variables so that the remaining variables do not make substantial improvements in the ability to predict a value of the dependent variables X_0 , given X_1, X_2, \dots, X_p . Therefore, (6) is rewritten as

$$X_{0j} = \alpha + \beta_i x_{ij} + \beta_2 x_{2j} + \dots + \beta_p x_{pj}$$

$$+ \beta_{p+1} x'_{p+1,j} + \dots + \beta_{k_n} x'_{k_n,j} + \epsilon_j \quad (8)$$

and an orthogonalization process is used to extract some of the k_n variables for use as predictors. The suggested procedure is as follows: Writing $x_{ij} = x_{ij}' - \bar{x}_i'$; $i = p+1, p+2, \dots, k_n$, define the matrix X with elements x_{ij} ; $i = 0, 1, 2, \dots, k_n$; $j = 1, 2, \dots, N$. The covariance matrix S is given by

$$S = (1/N) X X' \quad (9)$$

and the correlation matrix \hat{R} is defined by

$$\hat{R} = \left[\frac{s_{ij}}{(s_{ii} s_{jj})^{1/2}} \right]$$

$$i, j = 0, 1, 2, \dots, k_n \quad (10)$$

i.e., the element in the i th row, j th column, of \hat{R} is given by (10), where s_{ij} is the element in the i th row and the j th column of S .

Now define a family of $(k_n + 1) \times (k_n + 1)$ matrices $H_{ij}^{(1)}$, $j \neq 0$:

$$\left. \begin{aligned} h_{ii}^{(1)} &= 1 \\ h_{ij}^{(1)} &= -\frac{s_{ij}}{s_{jj}} \quad i \neq j \\ h_{ik}^{(1)} &= 0 \quad \text{otherwise} \end{aligned} \right\} \quad (11)$$

Then if $X^{(1)} = H_j^{(1)}X$, it follows that

$$\begin{aligned} x_{ik}^{(1)} &= \sum_{s=0}^{k_n} h_{is}^{(1)} x_{sk} \\ &= x_{ik} - \frac{s_{ij}}{s_{jj}} x_{jk} \quad i \neq j \\ &= x_{ik} \quad i = j \end{aligned} \quad (12)$$

This transforms X into a matrix $X^{(1)}$, whose rows give the residuals for each variable X_0, X_1, \dots, X_{k_n} after simple regression with X_j . Then, computing

$$\begin{aligned} S_j^{(1)} &= (1/N)H_j^{(1)}X X' H_j^{(1)'} \\ &= H_j^{(1)}S H_j^{(1)'} \end{aligned} \quad (13)$$

yields

$$\begin{aligned} Ns_{ik}^{(1)} &= \sum_{s=0}^{k_n} x_{is}^{(1)} x_{ks}^{(1)} \\ &= \sum_{s=0}^{k_n} \left(x_{is} - \frac{s_{ij}}{s_{jj}} x_{js} \right) \left(x_{ks} - \frac{s_{kj}}{s_{jj}} x_{js} \right) \\ &= \sum_{s=0}^{k_n} \left[x_{is} x_{ks} - \frac{s_{kj}}{s_{jj}} x_{is} x_{js} - \frac{s_{ij}}{s_{jj}} x_{js} x_{ks} \right. \\ &\quad \left. + \frac{s_{ij}s_{kj}}{s_{jj}^2} x_{js}^2 \right] = N \left[s_{ik} - \frac{s_{ij}s_{jk}}{s_{jj}} \right] \end{aligned}$$

in particular,

$$Ns_{00}^{(1)} = N \left[s_{00} - \frac{s_{0j}^2}{s_{jj}} \right] \quad (14)$$

which gives the residual sum of squares after regression of X_0 on X_j . More generally,

$$Ns_{ii}^{(1)} = N \left[s_{ii} - \frac{s_{ij}^2}{s_{jj}} \right] \quad i \neq j$$

gives the residual sum of squares after regression of X_i on X_j . In addition

$$\begin{aligned} Ns_{ij}^{(1)} &= N_{ij}^{(1)} = N \left[s_{ij} - \frac{s_{ij}s_{ji}}{s_{jj}} \right] = 0 \\ Ns_{ji}^{(1)} &= Ns_{ji} \end{aligned} \quad i \neq j$$

which establishes that $X_i^{(1)}$, the i th row vector of $X^{(1)}$ is orthogonal to $X_j^{(1)}$, the j th row vector of $X^{(1)}$ for all $i \neq j$. Finally, $s_{ik}^{(1)}$ gives the sample covariance of $X_i^{(1)}$ and $X_k^{(1)}$ for i and k .

Then the transformation from X to $X^{(1)}$ provides a new family of variables which are linear functions of the X 's and uncorrelated (orthogonal) with X_j . Essentially all effects due to X_j have been removed from all other variables. The corresponding correlation matrix is defined analogously with (10) as

$$\hat{R}_1 = \left[\frac{s_{ik}^{(1)}}{(s_{ii}^{(1)} s_{kk}^{(1)})^{1/2}} \right] \quad i \neq j, \quad k \neq j$$

If a second index $j_1 \neq j$ or 0 is now chosen and the same operations are repeated, i.e., let

$$X^{(2)} = H_{j_1}^{(2)} X^{(1)}$$

and

$$S_{j_1}^{(2)} = (1/N)H_{j_1}^{(2)} X^{(1)} X^{(1)'} H_{j_1}^{(2)'} \quad (15)$$

where $H_{j_1}^{(2)}$ coincides with (11), except for

$$h_{ij_1}^{(2)} = -\frac{s_{ij_1}^{(1)}}{s_{j_1 j_1}^{(1)}} \quad i \neq j_1$$

then a matrix $X^{(2)}$ is obtained whose rows give the residuals for each variable X_0, X_1, \dots, X_{k_n} after multiple regression with X_j and X_{j_1} , and a matrix $S_{j_1}^{(2)}$ for which $Ns_{00}^{(2)}$ is the residual sum of squares after regression of X_0 on X_j and X_{j_1} . All entries, except on the main diagonal with either a row or column index j or j_1 are zero and remaining elements give the sample covariances of $X_i^{(2)}$ and $X_k^{(2)}$.

In general, to define the $(m+1)$ th iteration choose j_m , an index, in such a way that $j_m \neq j, j_1, \dots$, or j_{m-1} and define

$$X^{(m+1)} = H_{j_m}^{(m+1)} X^{(m)} \quad (16)$$

and

$$\begin{aligned} S_{j_m}^{(m+1)} &= (1/N)H_{j_m}^{(m+1)} X^{(m)} X^{(m)'} H_{j_m}^{(m+1)'} \end{aligned} \quad (16)$$

where $H_{j_m}^{(m+1)}$ coincides with (11), except for

$$h_{ij_m}^{(m+1)} = -\frac{s_{ij_m}^{(m)}}{s_{j_m j_m}^{(m)}} \quad i \neq j_m$$

Then, continuing in this manner through k_n iterations, $Ns_{00}^{(k_n)}$ will be the residual sum of

of the multiple regression system and

$$1 - R^2 = \frac{s_{00}^{(k_n)}}{s_{00}}$$

is the multiple correlation coefficient.

Accordingly, a sequence of multiple correlation coefficients $R^{(1)}, R^{(2)}, \dots, R^{(k_n)}$, with $R^{(k_n)} = R$

$$1 - R^{(m)^2} = \frac{s_{00}^{(m)}}{s_{00}} \quad (17)$$

med.
ly,

$$R^{(1)} \leq R^{(2)} \leq \dots \leq R^{(k_n)} = R$$

Regression coefficients b_1, b_2, \dots, b_{k_n} can be obtained at each iteration where

$$s_{0j_{m-1}}^{(m)} \text{ for } m = 1, 2, \dots, k_n \quad (18)$$

where $j_0 = j$.

Therefore, the following multiple regression equation is obtained:

$$b_1 X_j + b_2 X_{j1}^{(1)} + b_3 X_{j2}^{(2)} + \dots + b_{k_n} X_{jk_{n-1}}^{(k_n-1)}$$

the independent variables

$$X_j, X_j^{(1)}, \dots, X_{jk_{n-1}}^{(k_n-1)}$$

orthogonal. By solving (15) for $X^{(m)}$ in terms of $X^{(m+1)}$, the usual multiple regression equation is obtained.

In particular, the orthogonality provides that b_1, \dots, b_{k_n} are stochastically independent and the respective sum of squares

$$= \Lambda[s_{00}^{(m-1)} - s_{00}^{(m)}] \quad m = 1, 2, \dots, k_n \quad (19)$$

independent and have the F distribution with $N - k_n - 1$ degrees of freedom. Therefore, equation (1) is obtained.

The appropriate F ratio for testing the significance of the sum of squares due to regression is

$$\frac{(SS_1 + SS_2 + \dots, SS_{k_n})(N - k_n - 1)}{k_n(SSE)} \quad (20)$$

In the regression analysis described above, the variable X_{j_1} was selected and its effects were

TABLE 1. Analysis of Variance

Due to	Degrees of Freedom
Total sum of squares	$N - 1$
Regression SS_1	1
Regression SS_2	1
.	.
.	.
Regression SS_{k_n}	1
Sum of squares due to regression $SS_1 + SS_2 + \dots + SS_{k_n}$	k_n
Residual SSE	$N - k_n - 1$

eliminated. A second variable, X_{j_2} , was then designated and eliminated, and so on. Thus the sums of squares $SS_1, SS_2, \dots, SS_{k_n}$ must be interpreted as follows.

SS_m is the further reduction in the sum of squares due to the m th variable after all earlier variables have been removed. Therefore it is noted that the order in which variables are eliminated will influence the magnitudes of sums of squares and the assignment of significance to independent variables. However, this is more than compensated for by the independence of the respective F ratios.

The F test for single variables then is

$$F = \frac{SS_m(N - k_n - 1)}{SSE} \quad (21)$$

and has the F distribution with 1 and $N - k_n - 1$ degrees of freedom. If this is found significant, it is asserted that $X_{j_{m-1}}$ is significant after $X_j, X_{j_1}, \dots, X_{j_{m-2}}$.

Finally, since it has been noted that the order in which variables are eliminated is relevant, a procedure is now examined for making these decisions. It was also indicated that it is frequently desirable to find a satisfactory prediction equation involving relatively few variables; hence it is also desired to exhibit a rule for prescribing when to terminate the process.

From (17)

$$[R^{(m+1)}]^2 - [R^{(m)}]^2 = \frac{s_{00}^{(m)} - s_{00}^{(m+1)}}{s_{00}} \quad m = 0, 1, \dots, k_n - 1 \quad (22)$$

The suggested rule then will be to choose X_{j_m} so as to maximize (22), where $R^{(0)} = 0$ by

definition. By virtue of the orthogonality this is equivalent to computing the simple correlations between all variables not eliminated previously and $X_0^{(m)}$ and choosing that variable as X_{i_m} whose simple correlation with $X_0^{(m)}$ is a maximum in absolute value. This may also be obtained directly from the correlation matrix R_m derived from the covariance matrix $S_{i_{m-1}}^{(m)}$.

The ratio

$$[s_{00}^{(m)} - s_{00}^{(m+1)}](N - k_n - 1)/s_{00}^{(k_n)}$$

has the F distribution with 1 and $N - k_n - 1$ degrees of freedom, and a reasonable rule is to continue so long as

$$[s_{00}^{(m)} - s_{00}^{(m+1)}](N - k_n - 1)/s_{00}^{(k_n)} \geq F_\gamma \quad (23)$$

and to stop otherwise. It is suggested that γ be chosen at a low level of significance, say $\gamma = 0.10$, or F_γ as the 10 per cent significance level of F , inasmuch as it may happen that X_{i_m} may not be significant but that some later variables still may be significant; at the worst this will require a few more iterations. This question will be discussed later. Equation 23, as stated above, cannot be readily applied, since $s_{00}^{(k_n)}$ is not available until all k_n variables have been eliminated; hence after each iteration a preliminary estimate of $s_{00}^{(k_n)}$ must be obtained for use in (23).

Let J_m be the set of variables not yet eliminated after m iterations. The rule (22) for determining X_{i_m} provides that $s_{00}^{(m+1)}$ will have been computed for all possible choices of X_i . Then S_m is defined as

$$S_m = \sum_{i \in J_m} [s_{00}^{(m)} - s_{00i}^{(m+1)}] \quad (24)$$

S_m gives the sum of the reductions in the residuals for each of the variables separately. Defining

$$\hat{s}_{00}^{(k_n)} = \begin{cases} 0, & \text{if } s_{00}^{(m)} - S_m < 0 \\ s_{00}^{(m)} - S_m, & \text{otherwise} \end{cases} \quad (25)$$

then

$$[s_{00}^{(m)} - s_{00}^{(m+1)}](N - k_n - 1)/\hat{s}_{00}^{(k_n)}$$

is computed and compared with F_γ .

Discussion of methodology. The above methodology does not insure derivation of an 'absolute best' estimating equation, but it should result in a close approximation to such an equation. The Taylor series expansion insures selection of exponential effects if these are better

than linear effects. It also insures selection of interactions between variables if warranted. These aspects are illustrated as follows.

The exact relationship of a soil moisture index to runoff, for instance, is not known. Assume indices of 1, 2, and 3 for dry to wet conditions. Is a medium condition, 2, twice as effective as a dry condition, 1, in producing runoff? Or is a wet condition, 3, three times as effective as a dry condition, or 1.5 times as effective as a medium condition? The Taylor series expansion to the third order would provide the indices shown in Table 2 from those assumed above.

TABLE 2. Indices of Soil Moisture

Condition	1st Order	2nd Order	3rd Order
Dry	1	1	1
Medium	2	4	8
Wet	3	9	27

The actual but unknown effects of antecedent soil moisture should be reflected by some indices between the extremes, 1 to 27, resulting from the expansion.

Similarly, the Taylor series expansion provides that all possible interactions (cross products) of the variables will be included in the analysis. The higher the order of expansion, the larger the combinations of such cross products will be.

The orthogonalization of the variables insures that each will be considered on its own merits after essentially all effects of the first selected variables are removed. In effect, in the orthogonalization process a first variable (that with the highest simple correlation with the dependent variable, X_0) is selected and its effects are removed from X_0 and the other independent variables. The residuals left after the effects of the first variable are removed are then correlated with the remaining independent variables, and the one having the highest correlation is selected and tested. After the second variable's effects are eliminated, a third is selected by the same process, and so on until no significant reduction is obtained in the residuals.

This orthogonalization, in effect, does mathematically what the graphical curvilinear multiple regression analysis method approximates. This latter method is described in many textbooks and will not be dwelt upon here. The

statement is true to the extent that the potential and interaction terms in the Taylor approximation approximate the true curvilinear relations in the several independent and the dependent variables. Such true relations could be closely determined by arbitrarily inserting exponential values in the analysis. For example, $P^{1.5}$ could be inserted if there were reason to believe this might be better than P^2 .

This technique also has several deficiencies. Differences in the order of removal of the variables, such as might be encountered as a result of sampling fluctuations, may produce regression equations with completely different coefficients and apparently very substantial differences in the assignment of statistical significance. These differences are largely apparent differences and are predominantly caused by the variables not being independent, but they present problems in the assignment of physical explanations to phenomena encountered.

The rule (22) for choosing the order of elimination of variables will have a tendency to magnify spurious effects if N is not large. The stopping rule is also not completely satisfactory since after several seemingly unimportant variables are eliminated, the next one may be highly significant. The rule, however, 'looks good' only one step in the process. No simple method has been found to evade this difficulty. Orthogonalization will not compensate for bad or poor data, nor will it correct the characteristic of hydrologic data to tend toward low variance with small events and large variance with large events.

TRIAL APPLICATION OF THE ORTHOGONAL TRANSFORMATION MODEL

The statistical model described herein was used in a trial analysis of the data from watershed 3-H at the Central Great Plains Experimental Watershed, Hastings, Nebraska. Watershed 3-H has a drainage area of 3.95 acres. The data used were obtained for 18 years during the periods 1939 to 1954 and 1958 to 1959. This watershed was chosen because it had a record-rain gage nearby, a standard rain gage on the edge of the watershed, good runoff records, a soil moisture sampling station adjacent to the watershed is roughly pear-shaped. It has an average land slope of 6.1 per cent and deep

loessial soils of silty texture. Land use was as shown in Table 3.

The general practice on the watershed was that land for corn (following wheat) was disked soon after April 15. The corn was planted in lister furrows about the middle of May and was usually cultivated three times, so that the soil was gradually leveled from the ridges to the furrows. The corn was husked in late fall and the stalks were left standing in the field. The land for oats (following corn) was double-disked about April 1, and the grain was drilled. The grain matured about July 1 and was cut with a binder, shocked, and threshed. The wheat land (following oats) was plowed about the middle of July and was disked and harrowed to control weeds and obtain a good seedbed. The grain was drilled about October 1. The wheat ma-

TABLE 3. Land Use and Tillage Practices on Watershed 3-H

Year	Precipitation, inches	Runoff, inches	Land Use and Tillage Practice
1939	13.91*	2.06*	Strip crop (corn & oats), contour
1940	12.19	0.62	Strip crop (corn & oats), contour
41	26.27	4.94	Corn, straight row
42	32.18	5.58	Oats, contour
43	15.88	3.19	Wheat, straight row
44	28.42	7.10	Corn, straight row
1945	20.02	3.31	Oats, straight row
46	26.32	5.40	Wheat, contour
47	18.73	3.37	Corn, straight row
48	16.65	2.11	Oats, straight row
49	26.69	8.68	Wheat, straight row
1950	21.20	6.26	Corn, straight row
51	34.57	11.84	Oats, straight row
52	22.10	5.71	Wheat, straight row
53	19.22	2.58	Corn, straight row
54	18.64	3.35	Oats, straight row
1955	14.41	No record	Wheat, straight row
56	12.91	No record	Sorghum, straight row
57	31.43	No record	Fallow, subtilled on contour
58	19.69	1.64	Wheat, mulch tillage
59	25.98†	6.51†	Sorghum, mulch tillage
Average	22.42‡	4.73‡	

* March through December.

† April through October.

‡ Omitting 1939, 1955, 1956, 1957, and 1959.

tured around July 1 the following year and was usually harvested with a combine.

Continuous records of rates of runoff (surface) were obtained from the watershed by means of a type H flume and water-level recorder. Rates and amounts of precipitation were obtained from the charts of the nearby recording rain gage. Soil moisture values to depths of 6 feet, by two 6-inch and five 12-inch increments, from the soil surface down, were obtained from samples taken twice a month from a location adjacent to the watershed and in the same field in which the watershed is located. Records of land use and cultural practices were available for the period covered by the runoff records. The variables used in the analysis were:

$X_0 = Q$ = the dependent variable, amount of runoff as measured at the flume for each separable hydrograph, in inches. Many events were multi-peaked hydrographs that resulted from variable rates and amounts of rainfall during a storm. When these compound hydrographs resulted from distinct periods of rainfall, they were separated into separate hydrographs by use of a typical recession curve drawn to isolate each storm. (The recession flows of a few runoff records were estimated in like manner. This was necessitated by silt in the stilling well on the flume preventing the recorder float from going down as stage receded.)

$X_1 = P$ = an independent variable, amount of precipitation, in inches, taken from the recording rain gage chart and associated with each separate runoff event. Also included were many rainfall events which caused no runoff. All rainstorms, regardless of size, were included if the Experiment Station staff had tabulated rainfall intensities. In addition, those rainstorms which were 0.1 inch or greater were included when clock-hour intensities could be read from the recorder charts. The storms which produced no runoff amounted to less than half of the total of 763 events.

$X_2 = I$ = an independent variable, rainfall intensity, in inches per clock hour. This time-interval intensity was used because it is the unit published by the U. S. Weather Bureau in hydrologic bulletins. The use of rainfall intensities for the maximum 60 minutes or for shorter time intervals, in river basin studies, would require obtaining microfilms of the original rain

gage charts of the stations operated by USWB. It was believed that this would be impractical for most planning and operational studies for water conservation and use projects.

$X_3 = SM$ = an independent variable, an index of soil moisture available for use by vegetation at the beginning of each rain. This index, in inches, reflected the amount of the top potential 1 inch of water which was available in the soil when the rainstorm or rain flurry began. This index was derived empirically from soil moisture depletion curves based on a study by *Thornton, Waite, Mather, and Carter* [1958], from the soil moisture records available for the field in which watershed 3-H was located, and from antecedent rains. This index ranged from near zero when conditions were dry to a maximum of 1.0 following antecedent rains. The upper limit, 1 inch, was one of definition since the index was selected to represent that portion of the uppermost 1 inch of available water.

Neither land use nor treatment indices was included in this trial application. It was desired to determine the over-all effects of the three uncontrollable variables, thus permitting the removal of their combined effects, prior to a study of such controllable variables as crops, treatment, and cultural operations. This further study remains for the future.

The trial application was performed as follows. The Taylor series expansion (5) was carried out to terms of third degree, and hence by (7) we have

$$k_3 = \binom{6}{3} = 20$$

Using (22), the variable first eliminated was: Precipitation \times Intensity \times Soil Moisture = $PISM = X_1 X_2 X_3$. The rule (22) required continuation, and $P^2 = X_1^2$ was next eliminated. Following this, $I \times SM = X_2 X_3$ was eliminated and then $P = X_1$ was eliminated.

Thus, a satisfactory regression model of the form

$$X_0 = a + b_1 X_1 X_2 X_3 + b_2 X_1^2 + b_3 X_2 X_3 + b_4 X_1$$

or

$$Q = a + b_1(PISM) + b_2(P^2) + b_3(ISM) + b_4(P)$$

was obtained using four of the twenty available variables. For this relation the unexplained

ity in X_0 was reduced 87.6 per cent. The summarized in Table 4.

ssion of trial applications. The residual squares used is not $Ns_{00}^{(k_n)}$ but rather since by terminating after four variables, is not available, but in the degrees of column $742 = N - k_n - 1$ degrees of were used. This results in underesti- the F ratio and is consequently con- e. The reduction of the degrees of is warranted since all k_n variables have ayed a role in the elimination procedure. ld be noted, in particular, that more tional methods would neither lead to this of variables nor indicate when to stop g additional variables. The four vari- selected by rule 22 have a coefficient of ination of 0.876, while all twenty vari- have a coefficient of 0.882. The four s selected by rule 22 each contributed owing to the coefficient of determination, e effects of the previously selected ones moved.

	Incremental Contribution to R^2
	0.81330
	.05286
	.00798
	.00167
	.00615

seen that with rule (22) variables may eted that are rather difficult to obtain nce subject to error, or variables difficult ain or understand functionally. Note that trial application the first variable was $\times SM$ and the third variable selected $\times SM$. Both these variables contain the ent soil moisture index—probably the eliable of the three parameters, since it imated—and clock-hour rainfall inten- his latter, while closely correlated with ime intensities, is probably not the inten- easure which is most closely associated noff from this small watershed. To avoid fficulty the computational process could rogrammed to remove first some arbi- selected variables such as P^2 or P , and o proceed with other selections based on (22).

ould be pointed out that a slightly differ- of data might result in a dissimilar re- m equation. This is due, in part, to the correlation between some of the variables.

TABLE 4. Analysis of Variance

Due to	Sum of Squares	df	F^*
Total Sum of Squares	54.90282	762	
Regressions			
<i>PISM</i>	44.65231	1	4859.48
P^2 after <i>PISM</i>	2.90241	1	315.87
<i>ISM</i> after P^2 , <i>PISM</i>	.43823	1	47.692
P after <i>ISM</i> , P^2 , <i>PISM</i>	.09185	1	9.996
Others		16	
Residual	6.81802	742	

* Significant at 0.01 level.

P and I , for instance, are highly correlated. *PISM*, P^2SM , and I^2SM , therefore, are different expressions of essentially the same thing. Only a minor change in some of the data could cause P^2SM or I^2SM to be chosen for first elimination. Similarly, *ISM* and *PSM* are much the same. This possibility would prevent application of the results of this analysis to other sets of data.

CONCLUSIONS

The methodology described herein permits the construction of a multiple regression equation in such a way that the awkward properties of the 'shotgun' method and evading difficulties associated with the experimenter's lack of information as to the appropriate choice of model are avoided. Under many conditions satisfactory results can be obtained with relatively few variables. However, no guarantee can be given that the resulting coefficient of determination will be a maximum for the number of variables actually used out of the k_n available variables.

The computational procedure's being iterative makes it readily adaptable to programming for automatic digital computers. The analysis of the Hastings data presented here was performed on the IBM 650 of the Bureau of Reclamation, U. S. Department of the Interior, Denver, Colorado.

In the trial discussed above, all twenty variables in the Taylor series were used. The effects of each expansion were not evaluated. It is believed that it would be desirable to test the gains in reducing the unexplained variance in Q by successively computing the coefficient of determination for the first-, second-, and third-order terms of the Taylor series. Should it be found, for instance, in samples of a long series

of analyses, that the coefficient of determination for the third-order term was not significantly greater than that for the second order, it would not be necessary to carry other expansions to the third order. This would reduce the number of variables to ten rather than to twenty. This might be advantageous in developing forecasting equations for snow-melt runoff in the West, for instance, for a whole series of river basins.

The statistical procedures outlined above, and as illustrated by its trial application, does not eliminate assumptions intrinsic in the multiple regression approach (see first paragraph of section on 'The Multiple Regression Model') but is believed to offer an approach adaptable to many hydrologic problems. Its chief virtue lies in our ability to evaluate with it the importance of many individual variables successively, after the effects of previously selected variables have been removed. In the ordinary multiple regression method the effect of each independent variable is inextricably mingled with the effects of each other independent variable, and the relative importance of selected ones cannot be determined.

APPENDIX

Index of Notations Having a Fixed Meaning

X_0 = dependent variable.

X_1, X_2, \dots, X_p = independent variables.

p = number of independent variables.

ϵ = independent, normally distributed, random variables with common variance.

$\beta_1, \beta_2, \dots, \beta_p$ = population regression coefficients.

α = intercept of population regression line.

b_i = sample regression coefficients.

a = intercept of sample regression line.

$X_{ij}; i = 0, 1, 2, \dots, p, j = 1, 2, \dots, N$ = j th measurement of i th variable.

N = sample size.

$\binom{n}{r} = \frac{n!}{r!(n-r)!}$ = binomial coefficient.

\bar{X}_i = arithmetic mean of i th variable.

$$\left[\frac{\partial^m}{\partial X_1 \partial X_2 \dots \partial X_m} f(X_1, X_2, \dots, X_p) \right]_{(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p)}$$

 = m th partial derivative of $f(X_1, X_2, \dots, X_p)$ with respect to X_1, X_2, \dots, X_m evaluated at $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p$.

R_n, R_n' = remainder term in Taylor series expansion.

k_n = number of variables introduced in Taylor series expansion to terms of order n .

A' = transpose of matrix A .

X = matrix of observations.

S = covariance matrix with elements $i, j = 0, 1, 2, \dots, k_n$.

\hat{R} = correlation matrix with elements $i, j = 0, 1, 2, \dots, k_n$.

$H_i^{(m)}$ = matrix employed in m th orthogonalizing transformation.

$S_{i, \dots, (m)}$ = derived covariance matrix after m th orthogonalizing transformation.

\hat{R}_m = derived correlation matrix after m th orthogonalizing transformation.

R = multiple correlation coefficient.

$R^{(m)}$ = multiple correlation coefficient using first m variables eliminated.

SS_m = reduction in sum of squares due to error as a result of eliminating m th variable.

SSR = sum of squares due to regression.

SSE = sum of squares due to error

F_γ = γ per cent significance level of F distribution.

Acknowledgments. Messrs. John Allis and Frank Dragoun, Agricultural Research Service engineers at the Great Plains Experimental Watershed at Hastings, Nebraska, supplied the data used in this analysis. Messrs. Perry Ford and Darrell Webber, U. S. Bureau of Reclamation, Denver, Colorado, programmed the statistical model and ran the data through the computer. Messrs. L. L. Kelly, Agricultural Research Service, H. O. Ogrosky, Soil Conservation Service, and H. S. Riesbol, U. S. Bureau of Reclamation, Chairman and members, respectively, of the Supervisory Group directing the activities of the Work Group of Cooperative Water Yield Procedures Study, advised and encouraged the author in this work and reviewed the manuscript. The development, and the analyses, could not have been completed without the able assistance of these people.

REFERENCES

- Sharp, A. L., A. E. Gibbs, W. J. Owen, and J. Harris, Application of the multiple regression approach in evaluating parameters affecting water yields of river basins, *J. Geophys. Res. search*, 65, 1273-1286, 1960.
- Thornthwaite, C. W., J. R. Mather, and D. Carter, Three water balance maps of eastern North America, Resources for the Future, Inc. 47 pp., November 1958.

(Manuscript received April 12, 1961; revised July 28, 1961.)

Changes in the Levels of Lakes Michigan and Huron¹

IVAN W. BRUNK

U. S. Weather Bureau, Chicago, Illinois

Abstract. Ten-year overlapping averages of mean annual levels of the Great Lakes for the period 1860 to 1960 indicate peaks in the levels of most of the lakes in the 1880's and around 1950. The level of Michigan-Huron averaged more than 1.5 ft lower in the 10-year period ending in 1955 than in a similar period ending in 1887. The apparent factors associated with this drop in levels are considered, and it is concluded that it is possible that natural and artificial changes in the natural outlet control system of Lake Huron have been responsible for practically all of this observed drop in Michigan-Huron levels in the 68-year period.

Introduction. The present form and connections of the Great Lakes are the results of a long series of geological events. Hough (1902) discussed the effects of the continental ice sheets which scoured and molded the land in the region several times in the last 10,000 years and indicated that the last period of glacial ice was about 11,000 'radiocarbon years' ago. He also pointed out that the level of Lake Michigan has been as much as 60 ft and 350 ft lower than the present level during the period since the retreat or melting of glacial ice. The most recent identifiable stage of glacial ice was substantially higher than the present level of Lake Michigan, approximately 580 ft above sea level was a stage of glacial ice, approximately 16 ft higher, about 2500 'radiocarbon years' ago, the Algoma stage. Since that time, when Hough believed a lateral shift of the outlet channel of Lake Huron onto more easily eroded material occurred, downcutting presumably continued without interruption until the present levels of Michigan-Huron² were attained. However, Hough also stated that there was no evidence of a change in M-H levels in the recent time and that it is generally assumed that the rate of downcutting at the present time is negligible.

The opinions and conclusions expressed in this paper are those of the author and not necessarily those of the U. S. Weather Bureau. 'M-H' will hereafter be used to designate Lakes Michigan-Huron, which have the same level because of the broad and deep connection through the Straits of Mackinac and are usually considered hydrologically as one lake.

Since 1860, systematic and continuous measurements of Great Lakes levels have been made by the U. S. Lake Survey. From lake level data furnished by that organization,³ Figure 1 was prepared. This depicts the long-term changes in the levels of the Great Lakes by means of 10-year overlapping averages. Two prominent features of Figure 1 are the peaks in levels in the 1880's and around 1950. On some of the lakes the early peak was comparable to the later one. However, the average level of M-H in the 10-year period ending in 1955 was significantly (1.61 ft) lower than the average for the 10-year period ending in 1887. In 1919 it was stated [Warren, 1921]: 'Whether or not there has been any marked change in the level of Lake Huron due to change in regimen of its outflow channel is still a mooted question, and probably will remain so unless the stages of the lakes should return to the high levels of the 80's.' It was the purpose of this study to learn why M-H did not return to the level of the 1880's, and in particular to see whether there has been any change in the regimen of its outflow channel.

Factors affecting M-H levels. The following factors could cause a significant change in M-H levels: (1) inflow from Lake Superior; (2) precipitation; (3) evaporation; (4) diversion;

³ All basic data concerning lake levels, outflow, and net basin supply (defined later), used in this study were furnished through the kind cooperation of the U. S. Lake Survey, Corps of Engineers, except that records of Lake St. Clair levels before 1898 are from the report of the International Waterways Commission, 1910.

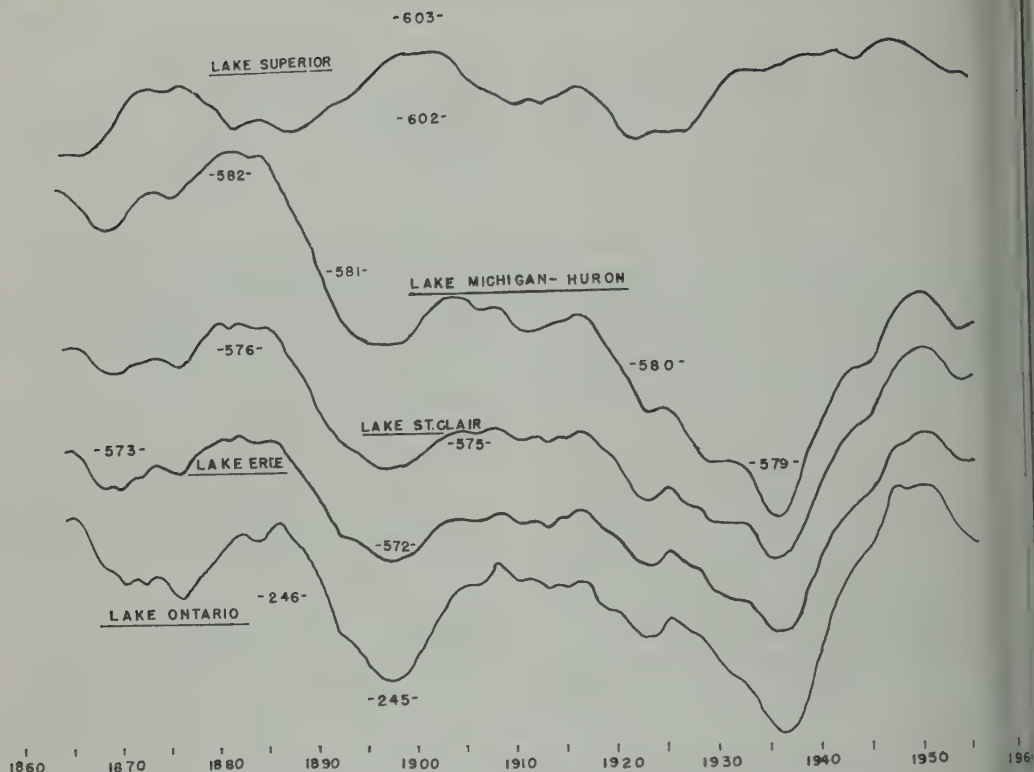


Fig. 1. Great Lakes levels, 1860 to 1960 (10-year overlapping averages, plotted at midpoint of 10-year period).

(5) crustal movements; (6) ice blockades; (7) changes in outlet conditions. Although all these factors affect the level and outflow of M-H, the first four do so solely through changes in water supply available to the outlet. As the water supply changes, the lake level and outflow adjust continually in the direction which tends to restore a balance between the water entering and the water leaving the lake in accordance with the stage-discharge relationship (for example, see Fig. 2). If no additional changes in water supply occurred after a given change, the lake would eventually reach a state of equilibrium in which the outflow was equal to the water supply, although considerable time would be required for adjustment to take place [*Senate Document 28, 1957*].

The last three factors listed change the lake level by a change in the stage-discharge relationship, as a result of a change in the physical characteristics of the outlet. Any change in these factors will result in a new curve, different from

but nearly parallel to the former curve. By a comparison of the stage-discharge relationships for the 10-year periods ending in 1887 and 1955, it will be shown that practically all of the observed drop of over 1.5 ft in M-H levels in that 68-year period was apparently due to the effects of changes in the natural outlet control systems of M-H. These changes, which resulted in the lowering of levels, will hereafter be referred to as 'downcutting.'

Evaluation of factors. Let us first consider crustal movements. The M-H lake level data used in this study are average levels referred to the Harbor Beach (Michigan) gage. The land at Harbor Beach is subsiding at a rate of 0.12 ft per 100 years with respect to the lake outlet [*U. S. Army, Corps of Engrs., 1952*], and thus after 100 years this effect alone would cause the gage at Harbor Beach to record levels that are 0.12 ft higher than the levels corresponding to given depths at the outlet. The effect of crustal movement between the 10-year period 1878 to

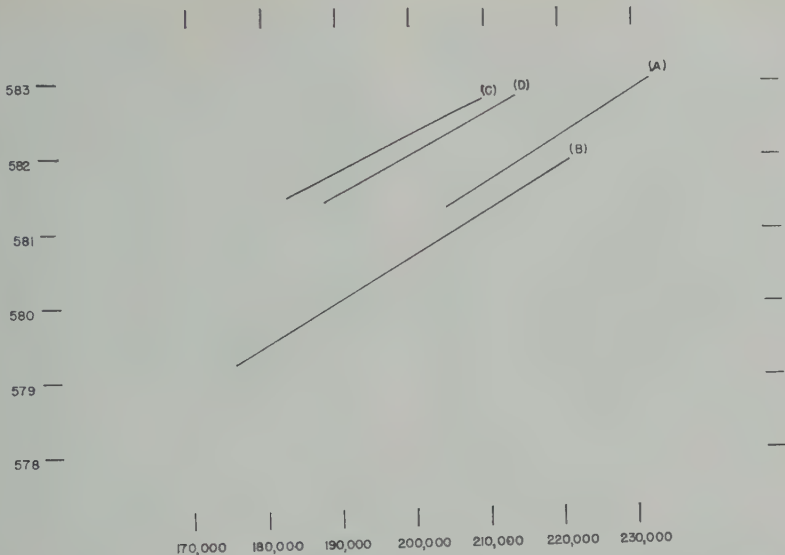


Fig. 2. Michigan-Huron stage-discharge relationships (discharge values corrected for effect of ice blockades: (A) 1878 to 1887; (B) 1946 to 1955, levels adjusted for effect of crustal movement from period (A) to (B); (C) 1878 to 1887, using derived discharge values; see text and Table 2; (D) 1878 to 1887, using 80 per cent of computed Erie net basin supply values to impute discharge values.

and the 10-year period 1946 to 1955 would cause a change of +0.08 ft in the level of

blockades which develop in the St. Clair Detroit rivers affect the level of M-H holding back large quantities of water at This keeps M-H continuously averaging 0.5 ft higher than if the rivers were free of ice [Freeman, 1926; U. S. Army, of Engrs., 1952]. Few specific data are available concerning the actual effect of ice obstructions in various years. To obtain an index of retardation of flow by ice, monthly values of level and its outflow through the St. Clair were plotted for the period 1860 to 1959. Estimate was made of the amount the flow reduced during each month when it appeared that the flow was less than could be

expected on the basis of a comparison with the normal stage-discharge relationship and the stage-discharge relationship for the preceding and following months. Adjustments were made as required for changes in elevation due to the water held back by the ice whenever this was significant. The average values of retardation of flow estimated as being due to the effect of ice blockades during the 100-year period 1860 to 1959 are indicated in Table 1.

The annual average retardation of 5754 ft³/sec is equivalent to an average additional elevation of 0.43 ft each year on M-H, as determined from the rating curve for this lake [Senate Document 28, 1957]. An approximation of the effect of ice blockades in each of the 10-year periods being compared can be made by considering the average yearly retarded flow, which

TABLE 1. Estimated Average Retardation of M-H Discharge Due to Effect of Ice, 1860 to 1959

	December	January	February	March	April	May	Annual
Retardation, ft ³ /sec	3,750	19,130	27,150	14,500	3,620	900	5,754
Retardation of total annual discharge (× 100)	5.5	27.7	39.3	21.0	5.2	1.3	100.0
Retardation as percentage of average flow, 1900 to 1959 (× 100)	2.1	12.5	18.8	9.0	2.0	0.5	3.2

was 4333 ft³/sec in the period 1878 to 1887 and 3783 ft³/sec in the period 1946 to 1955. The difference in the effect of ice in the second 10-year period as compared with the first would correspond to a change of -0.04 ft in the level of M-H.

Changes in outlet conditions. The discharge from M-H moves through the St. Clair River, Lake St. Clair, and the Detroit River to Lake Erie, a distance of about 84 miles, and there is no definite control section which determines the volume of outflow. Extensive improvements for navigation have been made in the control system, for under natural conditions the available draft was limited by the shallow water, only 2 to 6 ft deep, at the rim of the delta of Lake St. Clair [Freeman, 1926]. In 1961 a navigation channel 27 ft deep was available between Lakes Huron and Erie, and further deepening of channels was in progress.

In 1924 Horton and Grunsky [1927] stated that rating curves for the St. Clair and Detroit rivers had been materially changed by dredging to secure navigation depth and favorable channel alignment, and that much of this work was done in the period 1885 to 1897. Downcutting of the natural outlet control system of Lake Huron has also been attributed to a number of other factors, both natural and artificial, including: commercial removal of sand and gravel in the 16-year period after 1909 [U. S. Army, Corps of Engrs., 1952]; the natural erosion of the gravel bed in the Port Huron Rapids of the St. Clair River [St. Lawrence Seaway, Rept. of Engrs., 1952]; the effect of erosion or scour from the propellers of passing boats and the effect of ice gorges, which in one case was reported to have caused a deepening of several feet at places in the St. Clair Flat Canal [U. S. Chief of Engrs., Ann. Rept., 1904]. It has been stated [St. Lawrence Seaway, Rept. of Engrs., 1952] that the effect of changes in the St. Clair-Detroit River system between 1890 and 1925 resulted in the lowering of M-H levels by approximately 0.6 ft, and that any additional deepening of channels would further lower the levels [Jadwin, 1946]. Thus there has been considerable artificial downcutting in addition to any natural downcutting which may have occurred.

A comparison of the stage-discharge relationships for the two 10-year periods (ending in 1887 and 1955) will provide an estimate of the

effect of downcutting between the periods, adjustments are made for changes caused by ice blockades and crustal movements. The stage-discharge relationships for the periods 1878 to 1887 and 1946 to 1955, determined by the method of least squares from average annual values of M-H levels and discharge, are given as lines A and B, respectively, in Figure 2. The discharge values each year during both periods have been adjusted for the effects of ice blockades, and the levels during the second period have been corrected for the effect of crustal movement in the 68 years between the periods. The difference between the lines is about 0.4 ft, which is an estimate of the effect of downcutting between the periods. However, it will be shown that the discharge values during the first period are much too high and that line A should be replaced by line C, thereby indicating a much larger effect of downcutting.

A consideration of the water balance of Lake Erie provides a reasonably accurate method of checking the validity of the outflow of M-H, especially since approximately 90 per cent of the average net total supply to Erie (1900 to 1959) is provided by the outflow of M-H. Lake Erie levels have been affected but little by any changes in the regimen of outflow from that lake because the discharge through the Niagara River is controlled by a natural weir of rock. As Freeman [1926] indicated that the Niagara gagings were the most successful example of gaging the flow of a great river with extreme precision that has ever been made. Furthermore, obstructions by ice have very much less effect on the discharge of the Niagara River than on the outlet channel of M-H [Freeman, 1926].

The difference of the average yearly M-H outflow (not including diversion at Chicago) and total Erie outflow (including Welland Canal diversion) for the period 1900 to 1959 was over 22,000 ft³/sec, whereas for the period 1875 to 1899 it was less than 8000 ft³/sec. (The Erie basin would appear to have supplied only 31 per cent of the average yearly outflow of Erie during the earlier period, compared with approximately 10 per cent in the later period.) Since the difference in outflow is a close approximation of the water yield of the Erie basin (owing almost entirely to precipitation minus evaporation), the smaller value of average outflow in the earlier period is not reasonable.

of the higher rainfall in the Erie basin that period.⁴

discrepancy is also revealed in a comparison of the data for the net basin supply, or field, of the Erie basin. [Net basin supply = outflow minus inflow plus change in level (storage), or $NBS = O - I + S$.] The average yearly value of the *NBS* for the period 1900 to 1952 (no values available before 1952) was in excess of +2.5 ft, whereas for the period 1875 to 1899 it was less than +1 ft (both in feet on lake). These values are not different from the average yearly outflows from Lake Erie of approximately 22,000 and 8000 ft³/sec, respectively.

Since the *NBS* values are derived from observed or computed values of *O*, *I*, and *S*, the balance for *O*, *I*, and *S* with the values of *NBS* indicated above. But, since it has been shown [Freeman, 1926] that values of *O* for Lake Erie are apparently accurate,⁵ and since there is little reason to question the values of *S* obtained from records of lake level, the too low values of *NBS* in the early period indicate that the *I* for Erie for the early period is too low. The *I* for Erie is the *O* for M-H, and since the outflow records are derived from rating curves or formulas based on measured or assumed cross-sectional area of the outlet channel, the large values of M-H outflow before about 1890, therefore, would appear to indicate that the cross-sectional area of the outlet system was smaller than had been assumed before that time. In connection Horton and Grunsky [1927] pointed out that the determination of outflow from M-H prior to about 1898 requires special care since changes in outlet conditions have taken place more or less progressively.

Day's [1926] precipitation data, 1875 to 1924, were used in this study. Data for 1925 to 1952 were obtained by using as nearly as possible the same gauges, 21 in number, as were used by Day. [The precipitation in the Erie basin for the year 1888 was 30.62 in, instead of 40.62 indicated in Table 5.]

The correlation of Lake Erie levels and outflow (average annual values from 1860 to 1959), after corrections were applied to the levels for the effects of crustal movements, indicated a correlation coefficient of 0.97. Actually, the relationship between levels and outflow is curvilinear, the greatest departure from linearity being for low values of discharge.

To obtain a quantitative estimate of the correct outflow from M-H (inflow to Erie) before 1900, a regression equation was computed from the Erie data of 1900 to 1952 relating *NBS* and precipitation (correlation coefficient 0.83, normally explaining approximately 70 per cent of the variance). With this relationship, and the precipitation values before 1900 (no values available before 1875), new *NBS* values were computed. The average for the period 1875 to 1899 was over 24,000 ft³/sec, which is substantially different from the 8000 ft³/sec given above and much more reasonable when compared with the 22,000 ft³/sec for the post-1900 period. The new *NBS* values for the period 1878 to 1887 were used in the *NBS* equation (with unchanged values of *O* and *S*) to compute new values of *I* for Erie (Table 2). The stage-discharge relationship, with these new values of outflow used for M-H, yields line *C* of Figure 2. (The discharge values were adjusted for the effect of ice blockades in the various years of the 10-year period.) The difference between lines *B* and *C* is approximately 1.6 ft, which should be much nearer the effect of downcutting in the 68-year period than the value of 0.4 ft determined from line *A*. As indicated in column 7 of Table 2, the probable average annual discharge of M-H in the 10-year period 1878 to 1887 was 194,000 ft³/sec, compared with an annual average of 193,900 ft³/sec in the 10-year period 1946 to 1955. Therefore, the difference of 1.6 ft in average levels of the two periods could not have been due to any substantial difference in discharge.

It would be desirable to differentiate between the natural and artificial effects of downcutting. Considering the drop of approximately 16 ft in M-H levels in 2500 years, we would obtain an average rate of 0.44 ft in 68 years as the effect of natural factors between the two peak periods. (The level of Lake Erie has apparently not changed significantly in the 2500-year period, at least not in comparison with the 16-ft lowering of M-H levels.) It is likely that the rate would have been greater in the early part of the 2500-year period when the speed and volume of flow would have been substantially more, and before the present large delta had been built up in Lake St. Clair. (In 1915 it was estimated that the delta had filled about one-fourth of Lake St. Clair [U. S. Geol. Survey, 1915].) It would

TABLE 2. Computation of Probable M-H Discharge, 1878 to 1887
(All values in thousands of ft³/sec.)

Year	(1) Erie NBS	(2) Erie Out- flow	(3) M-H Out- flow	(4) (2) - (3)	(5) $S = NBS + I - O$ (1) + (3) - (2)	(6) Computed Erie NBS	(7) $I = O + S - NBS$ (2) + (5) - (6)	(8) Con- trolling factor
1878	+16	224	208	16	0	+42	182	+7
1879	+2	207	201	6	-4	+22	181	+4
1880	+5	212	204	8	-3	+30	179	+4
1881	+6	209	212	-3	+9	+30	188	+2
1882	+12	229	212	17	-5	+22	202	+3
1883	+9	224	220	4	+5	+30	199	+2
1884	0	226	221	5	-5	+20	201	+6
1885	+8	224	227	-3	+11	+29	206	+2
1886	-3	226	223	3	-6	+21	199	+10
1887	+4	224	218	6	-2	+19	203	+3
Average	+6	220.5	214.6	6	0	+26.5	194	+4

therefore be expected that the amount due to natural factors would have been less than 0.44 ft in the 68-year period studied, and it may even have been negligible, as was indicated by *Hough* [1958.]

Discussion. Because of the difference of approximately 1.6 ft indicated by the comparison of the stage-discharge relationships for M-H for a 68-year period, it appears that this change and resultant drop in M-H levels (for similar discharge values) can be attributed to natural and artificial changes in the outlet control system of Lake Huron. A more precise determination is beyond the scope and purpose of the study. However, there appears to be little doubt that a substantial change was associated with downcutting, since the effect of any other factors not included appears to be quite minor in comparison with the magnitude of the indicated change.

The principal limitation, in addition to the assumption concerning the accuracy of Erie discharge values, is the assumption that the relationship between Erie precipitation and *NBS* has not changed significantly with time. For example, it is possible that the water yield (or runoff) per inch of precipitation was less in the first 10-year period than in the second, because of the effects of deforestation and other changes. (However, there is no detectable time trend in the relationship between *NBS* and precipitation in the

Erie basin in the period 1900 to 1952.) But even if the water yield of the Erie basin in the first period were only 80 per cent of the computed *NBS* values for Erie, a value of approximately 1.4 ft would be indicated for the effect of downcutting in the 68-year period, as determined by the difference between lines *D* and *B* of Figure 2. However, it is by no means certain that the water yield was less in the early period, for *Horton and Grunsky* [1927] stated that cultural changes in the Lake Erie drainage basin had probably caused a *progressive* increase in water losses and a corresponding *reduction* in runoff.

Conclusion. A study of factors associated with the drop of more than 1.5 ft in M-H levels in a 68-year period has shown that very little could be accounted for except by a consideration of changes in outlet channel conditions. Although there may be reason to question certain aspects of the analysis and data employed, the results of the study indicate the possibility that natural and artificial downcutting of the outlet control system of Lake Huron has been responsible for practically all of this lowering of M-H levels.

Acknowledgments. The cooperation of the U. S. Lake Survey, Corps of Engineers, in providing the basic data concerning lake levels, outflow, and net basin supply, and the assistance of Lawrence A. Hughes, Research Forecaster, U. S. Weather Bureau, Chicago, Illinois, in the preparation of this manuscript, are gratefully acknowledged.

REFERENCES

- C., Precipitation in the drainage area of Great Lakes, 1875-1924, *Monthly Weather Review*, 4, 85-106, 1926.
- John R., *Regulation of Elevation and Level of the Great Lakes*, (A report to the Army District), Chicago, Illinois, 548 pp., 1927.
- Robert E., and C. E. Grunsky, Hydrology of Great Lakes, *Rept. Eng. Board of Review, District of Chicago, on the Lake Lowering Controversy and a Program of Remedial Measures*, part 3, appendix 2, Chicago, Illinois, 1, 1927.
- Jack L., *Geology of the Great Lakes*, University of Illinois Press, Urbana, 313 pp., 1958.
- Edgar, The Great Lakes, *Encyclopaedia Britannica*, vol. 10, 730-732, 1946.
- Document 28, 85th Congress, 1st Session, *Report of an Additional Diversion of Water from Lake Michigan at Chicago*, U. S. Government Printing Office, Washington, D. C., 74 pp., 1957.
- St. Lawrence Seaway, Report of the Joint Board of Engineers, U. S. Government Printing Office, Washington, D. C., 1927.
- U. S. Army, Corps of Engineers, *Preliminary Examination Rept. on Property Damage on the Great Lakes Resulting from Changes in Lake Levels*, appendix F, Levels of the Great Lakes, Great Lakes Division, Chicago, Illinois, 1952.
- U. S. Chief of Engineers, *Ann. Rept.*, 1904.
- U. S. Geological Survey, *Monograph 53, The Pleistocene of Indiana and Michigan and the History of the Great Lakes*, Government Printing Office, Washington, D. C., 1915.
- Warren, J. C., *The Diversion of Water from the Great Lakes and Niagara River*, containing report of W. F. Richmond, Assistant Engineer, House of Representatives, 66th Congress, Washington, D. C., 415 pp., 1921.

(Manuscript received June 3, 1961.)

Some Aspects of the Application of the Theory of Sediment Transportation to Engineering Problems

J. BOGÁRDI

*Hydraulics Laboratory, Research Institute for Water Resources Development
Budapest, Hungary*

Abstract. Sediment transportation in natural streams is to an appreciable extent influenced by hydrological conditions of the watercourse. Widely different hydrological conditions are encountered at various watercourses. Owing to these very differences, the theoretical laws governing sediment transportation that have been derived by considerations of hydrodynamics, or established on the basis of laboratory experiments carried out under more-or-less ideal conditions, can be used with the greatest conservatism only. Several examples of the effect of hydrological conditions on sediment transporting capacity, on the relationship between discharge and sediment concentration, and on the correlation between the particle diameter and mean velocity are listed in this paper.

Ordinary development has taken place in the theory of sediment transportation during the last decade. Although the results attained in the majority of cases are far from conclusive, the literature dealing with sediment movement has been enlarged by many valuable contributions. Many questions are yet to be answered and a number of the unsolved problems still encountered in the theory of turbulent water transport. Most of the available results have been obtained by laboratory research, which has been based on considerations of hydromechanics. Laboratory experiments carried out under more-or-less ideal conditions are highly useful for clearing substantial aspects of the phenomenon, yet these are the very causes of the difficulties encountered in practical application of the results. In fact, under natural conditions, sediment transportation never occurs in natural watercourses, or in extremely rare cases only, under the exclusive influence of hydraulic factors. Actually, besides the hydraulic factors, the regime of natural watercourses is influenced by a great number of circumstances, which can be termed collectively as 'hydrological factors'. These hydrological factors affect in the first order the formation of sediment in the catchment area, but may affect processes taking place in the river bed as well. A complete list of hydrological factors is practically unthinkable. However, to mention only the most important factors, the quantity and distribution of precipita-

tion, the geological structure of the catchment area, and the topography and vegetation of the latter belong in this category. The geological structure of the bed may also be regarded as one of the hydrological factors. Since sediment formation at times of winter precipitation differs appreciably from that in spring, summer, or autumn, the sequence of seasons may also be classified among the hydrological factors.

The effect of hydrological factors on sediment transportation of natural watercourses becomes manifest in extremely interesting and variable ways. In many instances even a correlative relationship can be established between sediment transportation and one or the other of the hydrological factors. It should be noted that a significant portion of these relationships is of the so-called symptomatic character, in which, as is well known, the factors involved are not interrelated, but each of them is in a causal relationship with the same additional factor or factors. Naturally, in the case of a given watercourse, this circumstance by no means affects the usefulness of the relationship.

A pertinent example of the effect of hydrological factors is the comparison of the sediment transporting capacity and the actual sediment load of watercourses. Sediment transporting capacity can be determined by applying the relevant methods of hydromechanics, of which several have been described in the literature. Sediment transporting capacity denotes the largest quan-

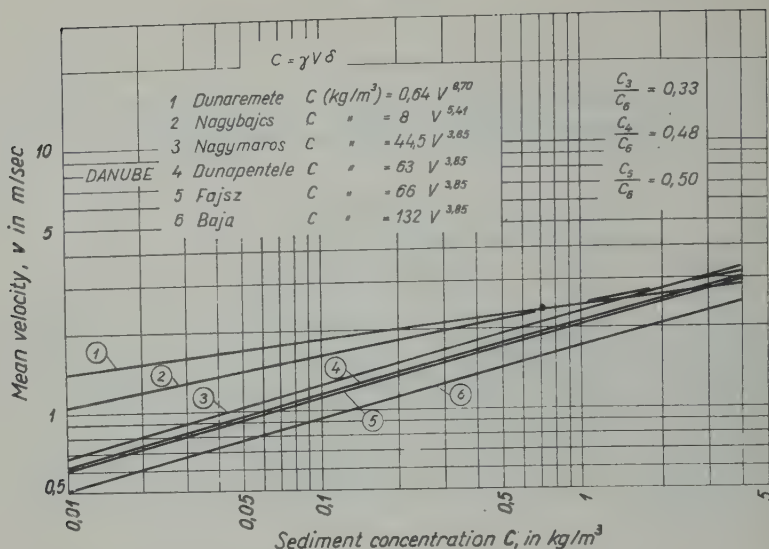


Fig. 1. Correlation between the sediment concentration and mean velocity on the Danube River.

tity of sediment that can be carried by the watercourse. The actual sediment load can at most attain this value, but is usually smaller, since, as indicated by experience, rivers seldom carry the amount of sediment that they are capable of transporting. This circumstance should naturally not be interpreted as an indication of discrepancy between the actual sediment transportation of natural watercourses and the laws of hydromechanics. Movement of sediment is always due to the energy of flow; yet, on the strength of the foregoing considerations, the rate of actual sediment transportation is usually governed by the availability of sediment material.

As established by investigations by the author, on the Danube River [Bogárdi, 1955a] and on many other rivers, when proceeding downstream the same sediment concentration C occurs at gradually decreasing mean velocities v . Figure 1 shows the relationship between the mean velocity v and sediment concentration C on the Danube River for six measuring stations based on field measurements. It has also been stated in the foregoing that watercourses carry generally less sediment than they are capable of transporting. An explanation is given thereby for the circumstance that, in cross sections lying farther upstream an appreciably higher mean velocity pertains to a given sediment concentration than in a cross section farther downstream. Obviously, the mean velocity corresponding to the momen-

tary discharge and wetted cross-sectional area as well as to roughness conditions, etc., is related to the available amount of sediment material. No account has, however, been given for the circumstance that in every cross section a higher sediment concentration pertains to higher velocities; i.e., the question why the sediment concentration C increases in the entire length of the river, in every cross section, according to an exponential function of v —as indicated for the Danube in Figure 1—remains unanswered. A certain explanation must be given for this circumstance since, as pointed out above, up to a certain limit the increase of sediment concentration does not necessarily involve—fr—from the hydromechanical point of view—higher velocities as well. The solution to the problem lies in the hydrological properties of the watercourse. Major precipitation giving rise to more intensive formation of sediment results in increased discharges, and these in turn necessitate higher velocities for their passage. The conclusion arrived at is that on the Danube, as obviously on many other rivers, sediment concentration increases with mean velocity according to a symptomatic relationship, since both sediment conditions and mean flow velocities are related to a third phenomenon, namely to the magnitude, distribution, and duration of precipitation falling on the catchment area.

Attention is called to Figure 2, which indicates



Fig. 2. Location of sediment observing stations along the rivers in Hungary.

ation of sediment observing stations
the Hungarian reach of the Danube and
r watercourses in Hungary.

connection with the relationships $C = f(v)$
d be noted further that these may not
d as functions in the mathematical
ut merely as stochastic ones relying on
ilities, inasmuch as the sediment con-
on C is obviously governed by a number
r factors besides the mean velocity v ,
are in a causal relationship partly with
d phenomenon referred to already, or
er phenomena. If, however, the relative
of particle fractions smaller than 0.01
sufficiently small—as in the case of most
urses in Hungary—then the $C = f(v)$ re-
ips can be established and may be
d and treated approximately as functional

manner similar to that used previously
ng the mean velocity v), sediment con-
on C can also be related to the discharge
corresponding expression, $C = f_1(Q)$, is
al to the foregoing one. These exponential
ships

$$C = dQ^l \quad (1)$$

etermined for most Hungarian water-
y by the author and were found to yield,
ral, acceptable results.

three rivers, the Danube, Tisza, and
Rivers, the number of observing stations

was greater than two (see Fig. 2), and this
offered the possibility for investigating the trend
of these relationships along the rivers. As
indicated by Figures 3, 4, and 5, the relationship
between C and Q hardly changes for stations
located along the same river [Bogárdi, 1955a].
More precisely, changes could be detected only
in the coefficient d of the exponential function.
The exponent l was found to remain constant
for the entire length of the river. It is obviously
the value of l that defines the rate at which the
sediment concentration C varies with changes
in discharge. As has been expounded in the fore-
going, the actual sediment concentration is in
most cases governed by the availability of
sediment material rather than by hydromechani-
cal factors, the former depending upon the
characteristics of the watercourse. Consequently,
the value of l is obviously controlled by the
character of the catchment area, by precipitation
conditions, etc., i.e., by the hydrological condi-
tions of the watercourse under consideration.

Pursuing this line of approach and taking into
account data furnished by 30 sediment observing
stations situated along 12 Hungarian rivers, the
author attempted to correlate the value of l to
the hydrological characteristics of the catchment
areas pertaining to individual sediment observing
stations [Bogárdi, 1956]. The relationship, which
is obviously of stochastic character, has been
determined by numerical and graphical correla-
tion. Among the hydrological characteristics

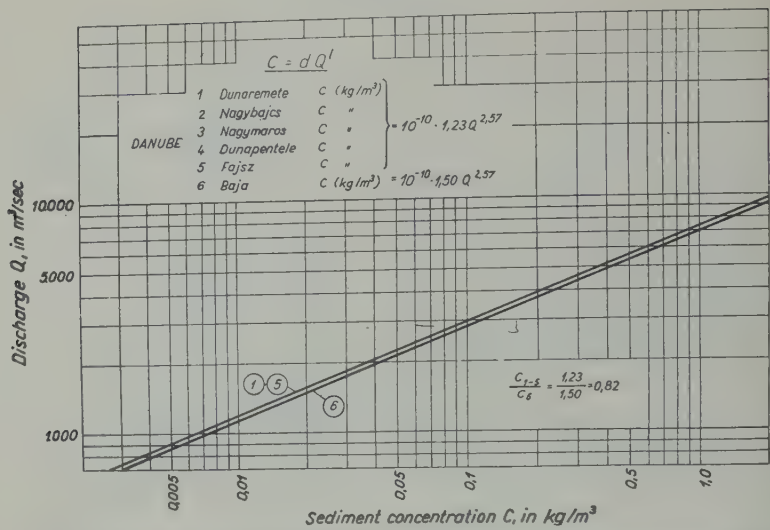


Fig. 3. Correlation between the sediment concentration and discharge on the Danube River.

four factors have been taken into consideration as the independent variables X_1 , X_2 , X_3 , and X_4 .

In these investigations

$$S = X_1$$

is the average width of the catchment area, which has been determined as the quotient of

the size of the catchment area and of the distance of the observing station from the origin of the river. The second variable

$$MQ = X_2$$

is the long-term mean discharge at the particular observing station on the river; the third,

$$HQ/NQ = X_3$$

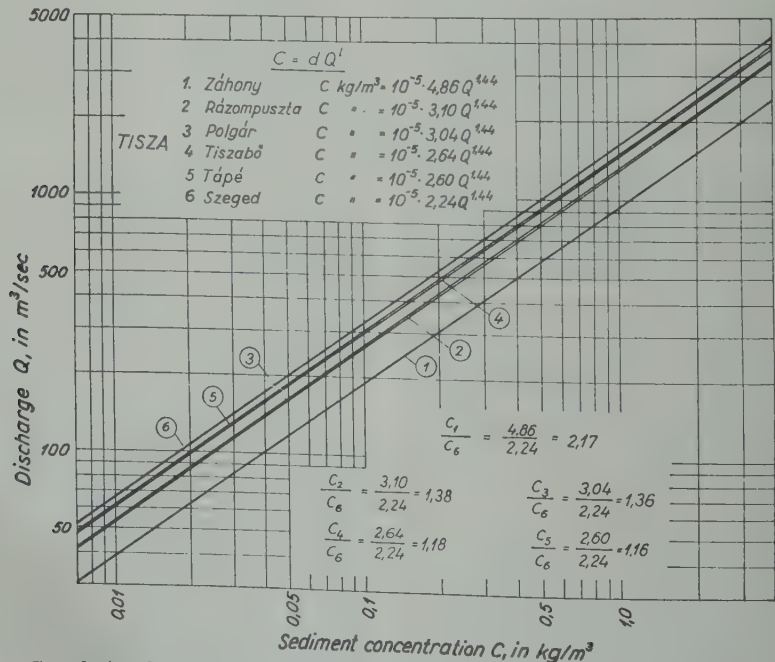


Fig. 4. Correlation between the sediment concentration and discharge on the Tisza River.

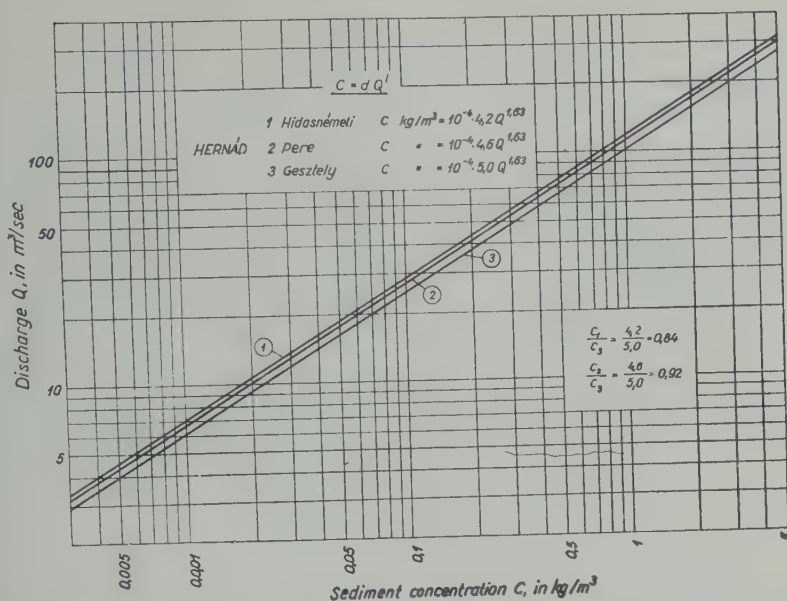


Fig. 5. Correlation between the sediment concentration and discharge on the Hernád River.

ratio of highest and lowest discharges,

$$M_q = X_4$$

specific mean runoff, which has been used as the quotient of the mean discharge and size of the catchment area.

Result of the graphical correlation is shown in Figure 6. As may be seen from the figure, a close relationship was found to exist between the dependent variable $l = Y$ and the mentioned four characteristic magnitudes. Investigation outlined above also supports the conclusion that the so-called hydrological characteristics of a watercourse influence in a decisive way the amount of sediment transported, such as it is the value of l which defines the amount of sediment concentration C which varies with discharge.

It is an established fact that, proceeding downstream, the particles of sediment transported by the watercourse generally diminish in size. This reduction in size can be observed most readily in the case of bed load transported along the bottom. The average particle diameter d of bed load trapped during observations and the corresponding mean velocities v in the cross section are shown in Figure 7 [Bogárdi, 1955b]. Downstream

along the river, the reduction in particle size is seen to follow a definite trend in accordance with the reduction of the mean velocity. Although the observation data scatter within a fairly wide range, the relationship

$$d = 0.15 + 0.47v^{8.13} \quad (2)$$

is sufficiently close and may be treated as a function as well.

Obviously, this relationship cannot be justified hydromechanically at all; neither is it related to the approximate quadratic relationship existing between the particle diameter d and the critical mean velocity. Yet no connection could be detected between the above expression and the relationships established in the theory of channel stability for the particle size d and the hydraulic factors of the watercourse (such as the gradient S , the waterdepth h , etc.) Thus Figure 7 obviously represents the widest variety of hydrological and morphological characteristics of the watercourse which affect the diameter of sediment particles transported by rolling along the bottom.

It is considered appropriate to attach an additional remark to the foregoing considerations. In general, no hydromechanical justification can be given for the relationships existing between the sediment transport of natural watercourses and the so-called hydrological

istics of the latter, i.e., for those among that can be derived. Since, however, all ships of similar nature, regardless of causal or symptomatic, rest on conditions actually prevailing in nature, they can be at variance with laws derived on the basis of hydromechanical considerations. In the case of mean velocity smaller than the critical velocity corresponding to the particle size consideration can pertain, according to Figure 7, to any selected particle diameter. Can any contradictions arise in connection with Figures 1, 3, 4, and 5. However, using this principle as a basis for checking, it should always be remembered that the figures do not relate to average values. The particle diameter d denotes the average diameter of particles trapped in the entire cross section, sediment concentration and mean velocity can similarly be understood as mean values relative to the entire section. It should be noted that on natural watercourses any relationship connected with sediment movement can be obtained today solely on the basis of average values. This circumstance is in many cases an insurmountable obstacle in the way of generalizing empirical relationships.

In conclusion, it can be established that the indirect effect of hydrological factors expressing properties of the watercourse is so involved that very few relationships can be generalized to include all watercourses.

However, when applying theoretical results to the case of sediment movement in watercourses, other circumstances arise which also require consideration. In numerous instances the data that would be necessary are inaccessible for measurement on natural watercourses. It suffices here to mention the difficulties encountered in determining the energy gradient, the surface slope, which difficulties are especially insurmountable on plain rivers and at reaches. There are similar obstacles to the determination of the bottom velocity, which has been reduced up to this very day in many respects to a theoretical notion only.

Very frequently the shortcomings of the theoretical relationships which present obstacles to their practical application. The role of these shortcomings during the determination of the relationship—usually under boundary conditions—is insignificant. Their

effects, being within the accuracy limit, are unobservable. The notion of critical mean velocity should be remembered here, which under largely identical experimental conditions—in particular at nearly identical water depths—defines with fair consistency the incipient stage of sediment movement. As far as mean velocities are concerned, this circumstance has long been known.

Drawing on investigations of H. K. Liu, M. L. Albertson, D. B. Simons, and E. V. Richardson, the author established the relationship between the particle diameter d and the channel stability factor introduced by himself [Bogárdi, 1943, 1955a]

$$b = d/hS \quad (3)$$

for various regimes of movement.

As revealed by these investigations, the different regimes of movement, including naturally the boundary condition between stagnation and incipient movement, are described by relationships of the form [Bogárdi, 1958]

$$b = d/hS = g d/U_*^2 = \beta d^{0.882} \quad (4)$$

where h is the water depth, i.e., the hydraulic radius, g is gravitational acceleration, S is the energy gradient, d is the particle diameter, and $U_* = \sqrt{ghS}$ is the shear velocity. In analogy to the shear velocity Reynolds number, the invariant quantity

$$U_*^2/gd$$

was termed by the author the shear velocity Froude number. Since, in keeping with the foregoing, the channel stability factor is

$$b = 1/\text{Fr}_* \quad (5)$$

consequently

$$1/\text{Fr}_* = \beta d^{0.882} \quad (6)$$

For various regimes, and for water having a temperature of 20°C and a sediment having a specific gravity $\gamma_1 = 2.65 \text{ g/cm}^3$, these relationships are shown in Figure 8. The thought that a direct relationship can be established between U_* and the particle diameter d is apparent from (4), whence

$$U_* = (g/\beta)^{1/2} d^{0.059} = \epsilon d^{0.059} \quad (7)$$

These relationships defining the various regimes

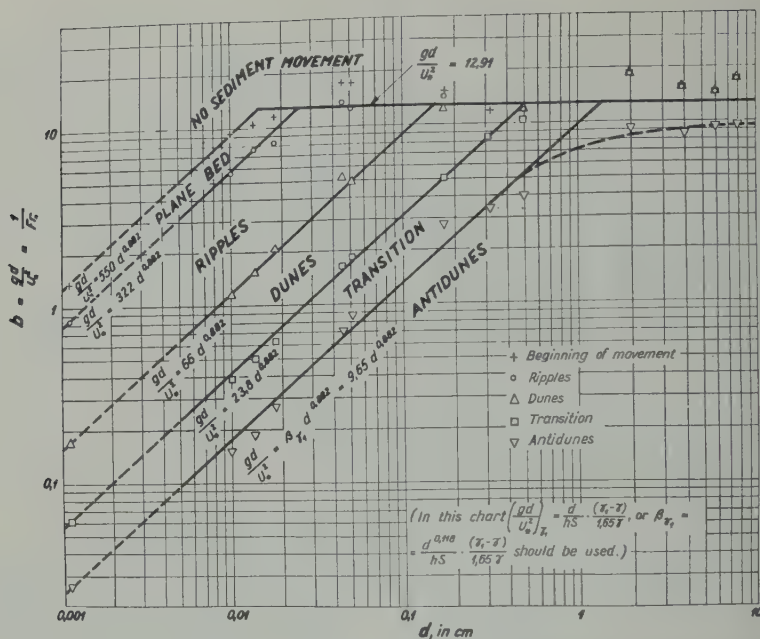


Fig. 8. Data of plane bed, ripple, dune, transition and antidune formation plotted in terms of $b = gd/U_*^2$ versus d .

of movement are shown again for the case of 20°C water temperature and $\gamma_1 = 2.65$ sediment specific gravity in Figure 9 [Bogárdi, 1959a, b].

It is clear, furthermore, that on the basis of either Figure 8 or Figure 9 a relationship can be developed between the critical tractive force τ_0 and the particle diameter d , since the tractive force

$$\tau = \gamma h S = \rho U_*^2 \quad (8)$$

The relationship is shown in Figure 10 [Bogárdi, 1959a, b], again for the case of 20°C water temperature and 2.65 sediment specific gravity. According to the figure, for $d < 0.0145$ cm

$$\tau_0(\text{g/cm}^2) = 0.00182d^{0.118} \quad (9)$$

and for $d > 0.0145$ cm

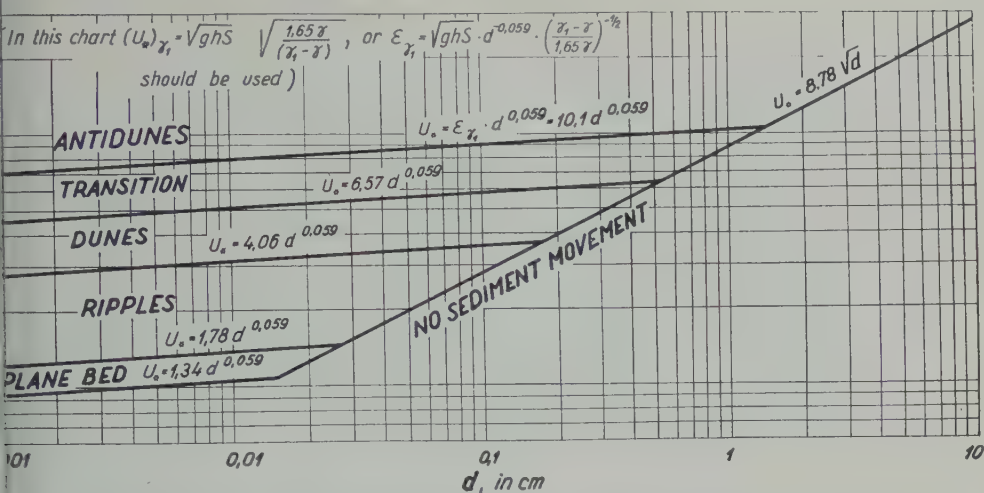
$$\tau_0(\text{g/cm}^2) = 0.0774d \quad (10)$$

It should be noted that the limit particle size $d = 0.0145$ cm shown in Figures 8, 9, and 10 is approximate since a certain zone of transition can be expected at the relationships represented in the three figures. Investigations are being carried out for clearing this aspect. For any given particle diameter d the regime of movement is defined, according to Figures 8

and 9, only by parameter β and parameter γ_1 respectively. The main advantage of the relationships thus is that sediment transport is defined by a single parameter.

No factor depending on either water temperature or specific gravity of the sediment is included in Figures 8, 9, and 10. This circumstance implies that these figures can be used regardless of the prevailing water temperature and specific gravity of the sediment.

It will be perceived, however, that if, under the influence of a change occurring either in the temperature of water or in the specific gravity of the sediment, the regime of movement pertaining to a given particle diameter d and the shear velocity U_* also suffers a change, then the above figures cannot be used unless the parameters are reduced in accordance with the changes mentioned. In fact, if the regime of movement is influenced by the temperature of water and by the specific gravity of the sediment, then obviously different limit values of the parameters will define the same regime of movement for every combination of temperature and specific gravity values. A possible approach to this problem would be to redraw Figures 8, 9, and 10 in accordance with the changed limit values of the parameters pertaining to different



9. Data of plane bed, ripple, dune, transition, and antidune formation plotted in terms of U_* versus d .

atures and specific gravities, but this be too laborious. Instead, the figures to a water temperature of 20°C and a t having a specific gravity of 2.65 can , but the parameters must be reduced to actual temperature and specific gravity

oaches based on different assumptions ade by the author in order to determine tent of temperature reduction [Bogárdi, 959a, b]. However, as revealed by recent gations, actual experimental results are a

prerequisite for the solution of the problem. Recent data further suggested that the original assumptions must be reconsidered, since the effect of water temperature is appreciably smaller. Very likely the effect of water temperature on regimes of sediment movement depends also on the particle diameter of the sediment. Therefore, the author is of the opinion that, since no reliable experimental data are as yet available, effects due to changes in temperature should, for the time being, preferably be neglected. As against earlier papers, reduc-

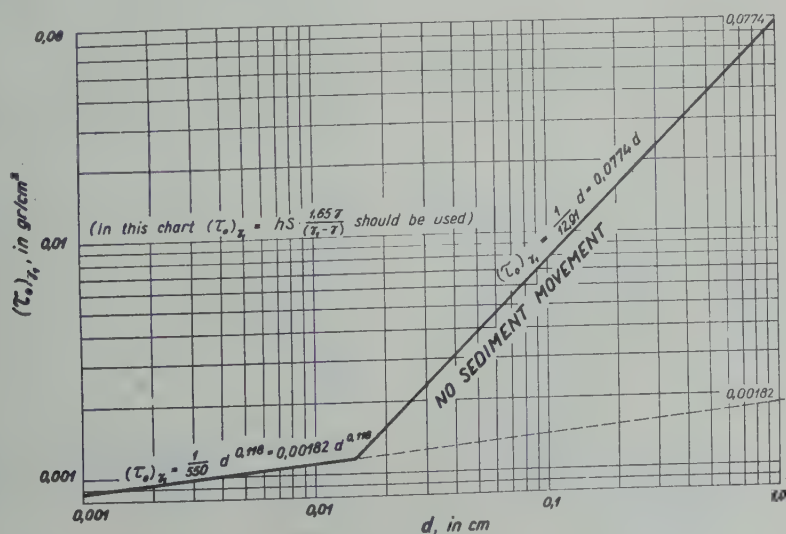


Fig. 10. Data of beginning of bed-load movement plotted in terms of $(\tau_*)_c$ versus d .

tions due to changes in temperature will be neglected in the following.

Effects of specific gravity of the sediment can be dealt with on a more reliable basis.

According to the frictional drag theory of sediment movement the critical tractive force is directly proportionate to the specific gravity of sediment in water, i.e., to a value of 1.65γ , if the actual specific gravity of the sediment is 2.65. This proportionality may justly be assumed to exist at regimes of movement other than critical. Relying on this assumption, if a tractive force τ_0' pertains to a certain regime of movement of sediment having a specific gravity of 2.65, then the same regime of movement will be caused by a tractive force

$$\tau_0[\gamma_1] = \tau_0'(\gamma_1 - \gamma)/1.65\gamma \quad (11)$$

for a sediment having a specific gravity γ_1 , which differs from the above value. The expression is in full agreement with practical experience, since for the same regime of movement a greater tractive force pertains to higher specific gravities of the sediment. Naturally, the actual tractive force is obtained from (11). If it is intended to use the relationships and figures relating to a specific gravity of the sediment of 2.65, the parameters must be reduced appropriately.

In Figure 10, and naturally also in equations 9 and 10, the reduction coefficient of the tractive force is

$$1.65\gamma/(\gamma_1 - \gamma)$$

i.e., Figure 10 and the aforementioned numerical expressions can be used for any specific gravity γ_1 by observing the value

$$\tau_{0\gamma_1} = hS1.65\gamma/(\gamma_1 - \gamma) \quad (12)$$

In Figure 8 the reduction for specific gravity is established in the following manner. Since the tractive force is proportionate to U^2 ,

$$\left(\frac{gd}{U_*^2}\right)_{\gamma_1} = \frac{d}{hS} \frac{(\gamma_1 - \gamma)}{1.65\gamma} \quad (13)$$

and

$$\begin{aligned} \beta_{\gamma_1} &= \frac{d}{hS} \frac{1}{d^{0.882}} \frac{(\gamma_1 - \gamma)}{1.65\gamma} \\ &= \frac{d^{0.118}}{hS} \frac{(\gamma_1 - \gamma)}{1.65\gamma} \end{aligned} \quad (14)$$

Naturally, it is sufficient to reduce only one of the values gd/U_*^2 and β .

In Figure 9 it again follows from the relationship of the tractive force and the shear velocity that

$$U_{*\gamma_1} = \sqrt{ghS} \sqrt{\frac{1.65\gamma}{(\gamma_1 - \gamma)}}$$

i.e.,

$$\epsilon_{\gamma_1} = \frac{\sqrt{ghS}}{d^{0.059}} \sqrt{\frac{1.65\gamma}{(\gamma_1 - \gamma)}}$$

Again, only one of the two values must be reduced.

As pointed out earlier, some uncertainty can be experienced in Figures 8, 9, and 10 around the particle diameter of 0.0145 cm. It has already been mentioned that further investigations are necessary to clear the effects of water temperature and specific gravity of the sediment. The relationships apply throughout to a single particle fraction only. Actually, however, mixtures of various particle sizes are always encountered. Quite obviously, much further research work is needed before the relationships shown in Figures 8, 9, and 10 can be applied to natural watercourses.

REFERENCES

- Bogárdi, J. L., *Bed Load Movement in River Regulation*, University Press, Budapest, 66 pp., 1958 (Hungarian).
- Bogárdi, J. L., *The Theory of Sediment Movement*, Hungarian Academy of Sciences, Budapest, 547 pp., 1955a (Hungarian).
- Bogárdi, J. L., Variation of bed load characteristics downstreams, *Rept. on the Activities of the Research Inst. for Water Resources Development in 1955*, 51-62, 1955b (Hungarian; English and Russian summaries).
- Bogárdi, J. L., Über die Zu- und Abnahme des Schwebstoffgehalts in den Flüssen mit Änderung des Abflusses, *Wasserwirtschaft*, 47(1), 59-66, 1956.
- Bogárdi, J. L., Some recent advances in the theory of sediment movement, *Hidrol. Közlemény*, 38(1), 241-252, 1958 (Hungarian; English and German summaries).
- Bogárdi, J. L., Channel stability and sediment movement, 6, *Convegno di Idraulica e Costituzioni Idrauliche*, A 17, Padova, 8 pp., 1959a.
- Bogárdi, J. L., Neuere Erkenntnisse auf dem Gebiet der Geschiebeforschung, *Österr. Wasserwirtschaft*, 11(12), 286-293, 1959b.

(Manuscript received April 25, 1961;
revised July 13, 1961.)

Sediment Transport of Glacier-Fed Streams in Alaska

WHITNEY M. BORLAND

*Sedimentation Section, Bureau of Reclamation
Commissioner's Office, Denver, Colorado*

Abstract. An empirical parameter is developed in terms of the glacier area, total drainage area, and length of watercourse. A graphical relationship is presented between the parameter, expressed as a fraction, and the sediment yield rate. Hydrologic records with particular emphasis on sediment data collected for glacier-fed streams provided the basis for the analysis leading to the development of the empirical parameter.

Throughout the literature on glacier geology there are numerous references to the large amounts of water released by glaciers and the amounts of sediment produced by them. Evidence of these facts is readily available. Although quantitative data are available, adequate data for planning and building engineering structures are lacking.

Examination of the data for active glaciers indicates that the runoff carrying a sediment load is more a function of temperature than of annual precipitation, which is the case. The relationship of suspended sediment to stream discharge (Fig. 2) is better developed for glacier-fed streams than for most other streams, probably because the suspended sediment near the glacier is relatively coarse. In most streams the fine-size sediments and silt are not as much a function of discharge as are the coarse-size sediments. The total sediment load of the streams decreases as the distance from the glaciers increases. This results in a gradation of the stream channel immediately below the snout of the glacier. The sediment material is coarsest next to the glacier and gradually decreases in size in a downstream direction. If a dam forms a reservoir which intercepts water onto the snout of the glacier, it is very difficult to estimate the amount of sediment deposited in the channel above the dam in addition to making a determination of the total sediment load at the dam site. The relationship of these two quantities is the sediment in the reservoir.

Studies made of the Susitna River Basin in central Alaska have resulted in confirma-

tion of the above general principles and have led to the development of a graph to estimate the total sediment yield rates for basins containing glaciers. In 1950 the U. S. Geological Survey began collecting sediment samples in several Alaskan streams. Hydrologic data collected by this agency [U. S. Geol. Survey, 1958, 1960] are for short periods and do not adequately cover the area. This is especially true of precipitation and temperature, for which there are no lengthy records of stations on or near the glaciers.

The roughly oval-shaped Susitna River Basin (Fig. 3) is about 19,900 square miles in area, being 250 miles long by 100 miles wide. On the west, north, and east, the Alaskan Range rims the basin, and on the east and south the Copper Plateau and the Chugach Mountains form its boundary. The mountains forming the rim of the basin, in general, rise to elevations of over 10,000 feet. A considerable portion of the central area of the basin is at altitudes ranging from 1800 to 3000 feet. It is covered with moss and low brush and contains innumerable lakes. Most of the larger tributaries of the Susitna River are fed by glaciers originating in the mountains of the Alaskan Range.

Recorded precipitation in the basin varies from 14 to 44 inches, but it is not unlikely that much higher rates of precipitation exist on the glaciers in the Alaskan Range, where the precipitation appears in the form of snow. The climate of the basin is continental; during the month of July the maximum average temperature is above 70°F and the daily minimum is about 44°F. Freeze-up of the river starts in early October in the higher regions, and in the

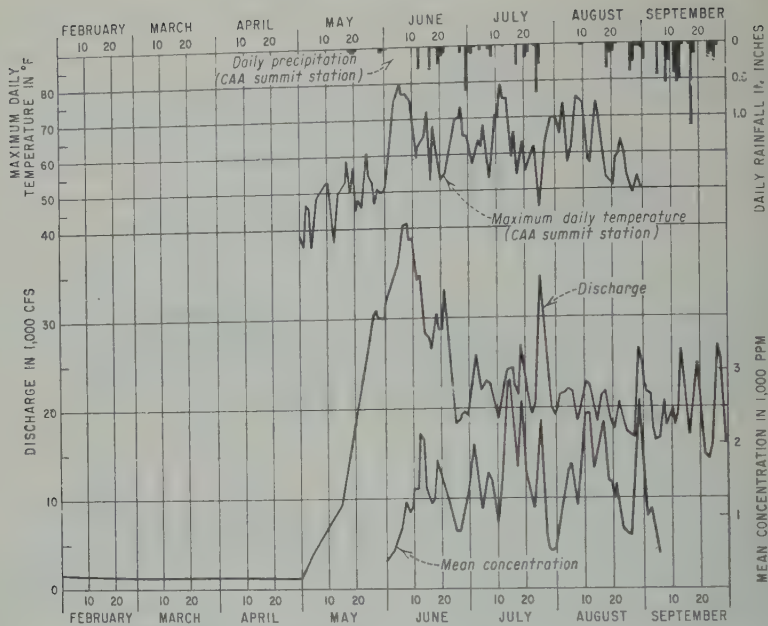


Fig. 1. Daily record of maximum temperature, discharge, rainfall, and mean concentration observed in 1957 for the Susitna River station at Gold Creek, Alaska.

spring most of the rivers are free of ice in late April or early May.

Stream-flow and sediment records are available for the Susitna River at Gold Creek and Denali. Because of turbulent flow, the Gold Creek station measures nearly the total load. At Denali total load measurements and computations have been made. Data pertinent to these two stations are listed in Table 1.

The principal tributaries between the two

stations are the Maclaren River, which is fed by glaciers, the Tyone River, which rises from lakes on the high plateau area, the Oshes River, and Kasine Creek. West Fork, Susitna, East Fork, and Maclaren are the principal tributaries feeding the river. For the sediment load at the downstream station to be less than the load at the upstream station, even with several tributaries contributing sediment in between, it is necessary that a portion of the sediment load

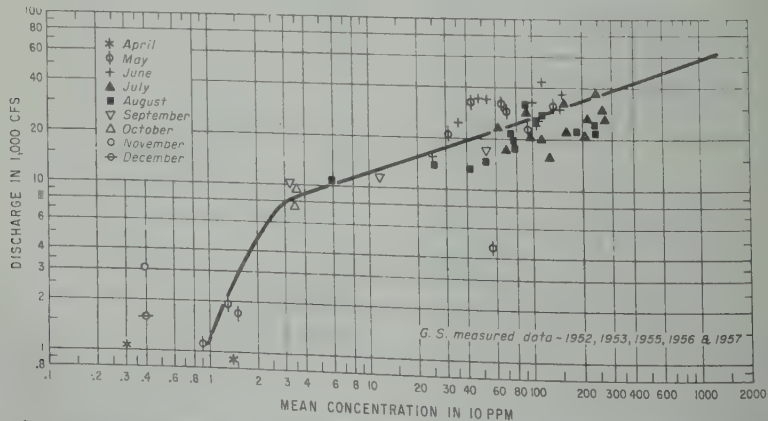


Fig. 2. Suspended sediment rating curve for the Susitna River station at Gold Creek, Alaska. The data were measured by the U. S. Geological Survey during the years 1952, 1953, 1955, 1956, and 1957.



Fig. 3. Location map showing the stream gaging and sediment sampling stations.

in the channel. Inspection of the river that rapid aggradation is taking place. the aggradation is so rapid that there ally no vegetation in the channel of the

akes and swamps form at the mouths tributaries that drain the high plateau y are formed because the bed of the eam, fed by glacier melt and sediment, faster rate than tributaries that are not aciers. A hydrograph of sediment load off (Fig. 1) indicated that practically ediment in the stream is moved when ers are melting. Also, the high plateau h its cover of low brush and moss and ined drainage pattern, would yield small of sediment.

A parameter, λ , was empirically defined as follows:

$$\lambda = (A_T/A_G)l$$

where

- A_T = total drainage area in square miles above the sediment sampling station.
- A_G = glacial drainage area in square miles.
- l = length (in miles) of the water course as measured from the snout of the glacier to the sediment sampling station.

This parameter was computed for each of the five sampling stations shown in Figure 4 and related to the corresponding annual sediment yield rate determined for the station. As noted from this graph a well-defined relationship is

TABLE 1. Comparative Data for Denali and Gold Creek Stations

Drainage Area, mi ²	Average Annual Runoff, acre-feet	Glacier Area, mi ²	Distance below Snout of Nearest Glacier, miles	Average Annual Total Sediment Load, acre-feet	Annual Total Sediment Yield Rate, acre-ft/mi ²
868	2,500,000	180	20.9	9120	10.50
6290	7,000,000	215	156.2	6440	1.02

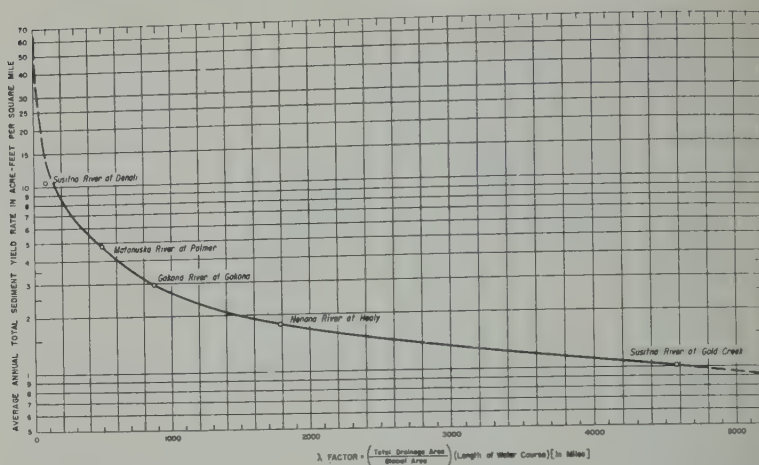


Fig. 4. Average annual sediment yield rate versus the empirically developed λ factor.

indicated for the few records gathered so far. A study of other Alaskan glaciers resulted in a sediment yield rate of approximately 65 acre-feet/mile²/year for the conditions when $A_G = A_T$ and $l = \lambda = 0$. From this preliminary study, the curve in the range above a yield rate of 10 acre-feet/mile²/year can be extrapolated to intersect this 65-yield rate as a limiting value indicated in Figure 4.

The slope of the stream immediately below the glaciers varies from about 20 to 50 feet/mile. A short distance below the glaciers, the slope of the stream flattens out. The slope of the Susitna River, from below the Denali gaging station to near the glacier, averages 8.5 feet/mile. The annual total sediment load at the snout of glaciers is estimated to be about 12,000 acre-feet and the load at the Denali station is 9120 acre-feet. Considering the width and length of the stream course, this difference of 3000 acre-feet results in an average deposition of about 1 inch/year.

Three other glacier-fed streams have been plotted in Figure 4, the Nenana near Healy, the Matanuska at Palmer, and the Kokona River at Kokona. The Nenana and Kokona rivers rise in

the Alaskan Range, and the Matanuska rises in the Chugach (Coastal) Range. The correlation shown in Figure 4 appears valid for the southern central portion of Alaska. There is considerable difference between the precipitation rates on the Matanuska glacier and on the glaciers in the Alaskan Range. Precipitation rates are over 100 inches/year on the Chugach Range. Despite this difference, however, the relationship still seems to hold for the Matanuska glacier, with widely different climatic conditions. It is not known whether this relationship or a similar one would apply to glacier-fed streams outside southern central Alaska. These data are presented for guidance only, with the hope that others may have information that can be added.

REFERENCES

- U. S. Geological Survey, Quantity and quality of surface waters of Alaska, *Water Supply Paper 1466*, Oct. 1950 to Sept. 1953, 243 pp., 1958.
- U. S. Geological Survey, Quantity and quality of surface waters of Alaska, *Water Supply Paper 1570*, 1958, 120 pp., 1960.

(Manuscript received February 3, 1961; revised June 23, 1961.)

Variation of Soil Erodibility with Geology, Geographic Zone, Elevation, and Vegetation Type in Northern California Wildlands

J. E. ANDRÉ AND H. W. ANDERSON

*U. S. Department of Agriculture, Forest Service
Pacific Southwest Forest and Range Experiment Station
Berkeley 1, California*

Abstract. Samples of the surface 6 inches of mineral soil were taken at 168 places in northern California and analyzed for the physical characteristics which index erodibility of the soil. Samples were selected in the major soil-geologic types of California, under standard conditions of slope (west, 20 per cent), at 1000-foot intervals of elevation (1000 to 4000 feet), by vegetation types (forest, brush, and grass), and in three separate zones (North Coast, Central, and Sierra). A multiple regression analysis related the 'surface-aggregation' and 'dispersion' ratios, as the indexes of erodibility, to geologic type, vegetation type, zone, and elevation and to their interactions. The surface-aggregation ratio was somewhat more significantly related to soil erodibility than was the dispersion ratio. Soil developed from acid igneous rock about $2\frac{1}{2}$ times as erodible as soil developed on basalt. Erodibility was highest for soils in brush, next under trees, and least under grass. No clear-cut relation of erodibility to elevation was found. The interaction of zone and geologic rock type showed significant variation in erodibility. The prediction equation explains 52 per cent of the variability in erodibility ratios. By combining predicted erodibility from this equation with chemical base status, for a sample of 20 of the soils, the explained variance was improved. Application of these relationships in studies of sedimentation from watersheds is illustrated.

INTRODUCTION

can menace water supply by creating water, injuring stream channels, and filling reservoirs. Some soils erode readily and do not. This paper reports a study of erodibility of California wildland soils—a study in evaluation of local sedimentation and their control.

Can we evaluate the erodibility of wildlands of a region when the soils are largely unknown and their characteristics unknown? If we have a soils map, we appeal to geologic, geologic, and topographic maps; for soil characteristics we sample soils having different characteristics and determine the erodibility characteristics in the laboratory and then use soil erodibility characteristics to map erodibility by regression analysis.

In western Oregon [Anderson, 1954] it was found that soil characteristics associated with different rock types were related to measured sediment discharge from watersheds. Jenny suggested that any soil characteristic

is the function of the soil-forming factors of the parent material (rock type in most mountain soils), climate, relief, organisms (largely vegetation), and time.

This study extends the studies of soil erodibility to three zones in northern California and extends the study of factors affecting erodibility to include vegetation type and climate, as indexed by geographic zone and elevation. Briefly explored, too, is the possibility of combining physical map characteristics and chemical base status of the surface soils in estimating erodibility.

PAST WORK

Classic work on the erodibility of soil was done by Middleton [1930]. He analyzed samples of several erodible and nonerodible soils, testing their physical and chemical properties. He found that none of the chemical properties studied could be used to differentiate between erodible and nonerodible soils; however, some physical properties of the soil could be so used.

The dispersion ratio, colloid to moisture equivalent ratio, and the erosion ratio were found to be valuable in separating erodible from non-erodible soils.

Erodibility has been considered a function of various soil properties such as permeability, texture, structure, and ease of dispersion. *Baver* [1932] expressed a belief that erodibility varies directly with ease of dispersion and inversely with permeability, aggregation, and particle size. *Lutz* [1934] and *Diseker and Yoder* [1936] studied available data and reported that permeability and ease of dispersion are the foremost factors in erodibility. They concluded that one of the principal differences between erodible and nonerodible soils is the degree of aggregation of the finer mechanical separates into large, stable granules. They found that eroded particles were aggregates rather than mechanical separates, showing the importance of knowing not only the amount but also the size distribution of the soil aggregates. A method of determining the amount of water-stable aggregates was devised by *Yoder* [1936]. His wet sieve analysis correlated well with erodibility as observed in the field.

In a flume study in which different soils were used on a bottom section, *Smerdon and Beasley* [1959] found a close relation between amount of erosion and the dispersion ratio. In a study of erosion in southern California, *Anderson* [1951] tested the relationship of several physical characteristics of the soil to measured suspended-sediment discharge from watersheds; this work demonstrated the usefulness of erodibility indexes, and Anderson recommended Middleton's dispersion ratio because of its simplicity and usefulness.

In a similar study of erosion from the watersheds of western Oregon, *Anderson* [1954] hypothesized that erodibility of a soil depended on the surface of the soil requiring binding—fine sand size and larger—versus the binding quality of the clays. He introduced a new index of soil erodibility—the surface-aggregation ratio. This index was highly correlated with suspended-sediment discharge from 33 watersheds when used in a multiple regression analysis. The relationships obtained indicated relative erodibility of the soils developed on different geologic rock types and permitted a prediction

of sediment yield to be expected with change in selected variables.

Recently *Wallis and Stevan* [1961] took of our northwestern California soils for study. They determined the relation of our erodibility ratios to their measurements of the chemical base status of the soils. They found that the dispersion ratio and the surface-aggregation ratio were highly correlated with the Ca + Mg adsorbed on the soil clays.

METHODS

Soil sampling. Relative erodibility of the surface soils was obtained by taking all soil samples under standardized conditions in which one geologic rock type, vegetation, elevation, and aspect zone were different. This method [*Anderson* 1954] requires that several soil samples of each type be taken under constant conditions of topography and vegetation condition; from these samples is determined what *Jenny* [1941] called a sequence. Soil samples were taken from the 0- to 6-inch depth of soil, at constant slope (20 to 30 per cent) and aspect (west), and under full natural vegetation cover, with an attempt being made to obtain a good distribution among geologic parent materials, elevations, and geographic zones (Table 1).

In California, the primary sampling was to determine a lithosequence; hence, geologic types in three geographic zones were sampled. Additional samples were taken to determine a climatic sequence (elevation) and a vegetation sequence (vegetation type). These sequences were expected to yield valuable information in interpreting sources of variation in soil erodibility.

Details of the soil sampling plan follow:

1. *Geologic types (lithosequence).* The geologic types, shown in the 1938 State Geological map by Jenkins and in the 1955 North Coast Map by Irwin and Tatlock, were grouped into 8 major rock types expected to differ in soil erodibility.

Igneous	Metamorphic	Sedimentary
Acid	Schist	Alluvium
Basic	Other metamorphics	Soft sediment
Ultrabasic		Hard sediment

Soft sediments were defined as sedimentary rock deposited during the Miocene period and

TABLE 1. Soil Sampling Pattern for Erodibility Determination, Northern California, 1958

Elev., ft	Acid Igneous			Basic Igneous			Serpentine			Other Meta.			Schist			Hard Sed.*			Soft Sed.†			Alluvium		
	F	B	G	F	B	G	F	B	G	F	B	G	F	B	G	F	B	G	F	B	G	F	B	G
1000																								
2000																								
3000																								
4000																								
1000																								
2000																								
3000																								
4000																								
1000																								
2000																								
3000																								
4000																								
1000																								
2000																								
3000																								
4000																								
1000																								
2000																								
3000																								
4000																								

* Fine sediments older than Miocene.

† Fine sediments Miocene and younger.

F, B, and G are forest, brush, and grass vegetation types.

Hard sediments, those from deposits before the Miocene period.

Climatic sequence. A climatic sequence was established by taking samples at four elevations, 1000, 2000, 3000, and 4000 feet, and in separated areas in the three geographic regions, North Coast, Central Coast, and Sierra Nevada. Thus, samples were taken in areas with annual rainfall ranging from 35 to 50 inches and mean annual temperatures ranging from 55°F. Microclimatic differences were expected to be minimized by taking all samples under standard conditions of 20 to 30 per cent relative humidity, west exposure, and full vegetation cover.

Vegetation sequence. Additional soil samples were taken at each vegetation type on each geologic rock types and elevation levels, and on each brush type on acid igneous rock type at 1000, 2000, 3000, and 4000 feet. These additional samples permitted evaluation of the influence of vegetation on soil erodibility properties and of the interaction of vegetation with soil type. In all, 168 soil samples were taken (Table 1).

Analysis of the soil samples. The soil samples were analyzed in the laboratory for texture

and Middleton's suspension characteristics. Soil texture was determined with sieves and hydrometer; Middleton's suspension percentage was also determined with a hydrometer. Details of soil sampling locations and results of the individual soil analyses are available upon request [André, 1960]. Various erodibility indexes were computed from the laboratory textural and suspension measurements.

Indexes of erodibility. Two indexes of soil erodibility were determined, the surface-aggregation ratio and the dispersion ratio. Specific definitions of the soil characteristics going into the indexes and of the indexes themselves are:

Surface. Amount of surface in cm²/g on particles larger than silt (larger than 50 microns in diameter); soil particles were considered to be spheres and to have a density of 2.65.

Aggregated silt and clay. Total percentage of silt and clay in dispersed soil minus percentage of that in an undispersed soil. The latter is the suspension percentage, determined by the method of Middleton [1930; see Anderson, 1954].

Surface-aggregation ratio. Surface divided by aggregated silt and clay—the ratio of the

TABLE 2. Means and Standard Deviations of Surface-Aggregation Ratios and Dispersion Ratios for the Different Rock Types in Northern California

Item	Rock								A Co bind
	Acid Ign.	Basic Ign.	Serp.	Misc. Meta- mor.	Schist	Hard Sed.*	Soft Sed.†	Allu- vium	
Number of samples	38	19	19	23	14	16	28	11	10
Mean surface- aggregation ratio	118	49	41	46	89	61	78	124	7
Standard deviation, surface-aggregation ratio	41	26	18	23	60	11	65	106	5
Mean dispersion ratio	61	53	41	51	63	47	49	51	5
Standard deviation, dispersion ratio	10	9	13	8	13	7	12	16	1

* Marine sediments older than Miocene.

† Marine sediments Miocene and younger.

amount of surface 'requiring binding' to the amount of 'binding' clay present in the soil [Anderson, 1954].

Dispersion ratio. Suspension percentage divided by total silt and clay [Middleton, 1930].

Tests of erodibility indexes. Means and standard deviation of the soil characteristics and erodibility indexes are summarized by rock types in Table 2. The indexes were first tested for their predictability from the soil-forming factors: geologic rock type, climate indicated by zone and elevation, and vegetation type. Multiple regression analyses, without interactions, were made in this first test of the 168 soil sample results.

Partial regression coefficients of the soil characteristics and the erodibility indexes with geologic rock type, cover, zone, and elevation, together with tests of significance of the erodibility indexes and their interactions, are given in Table 3. Dispersion ratio, surface-aggregation ratio, surface, and aggregated silt plus clay had multiple correlation coefficients of 0.70, 0.72, 0.78, and 0.68, respectively. The standard error of estimates are 9.1, 11.9, 5.5, and 7.4, respectively. In the remainder of this paper only the two indexes of erodibility, surface-aggregation ratio and dispersion ratio, are discussed.

Relation of surface-aggregation ratio to geology, zone, elevation, and vegetation. The relation of soil erodibility—the surface-aggre-

gation ratio S/A —to the geologic type, zone, elevation, and vegetation type and their interactions are given in Table 4. The results are in the form of a six-variable prediction equation. For a particular combination of conditions, the surface-aggregation ratio (S/A) can be estimated, being equal to a constant, 77.38, plus the effect of geology (G) plus the effect of cover (C) plus the effect of elevation (E) plus the effect of zone (Z) plus the interaction effect of $G \times Z$ and $G \times C$.

These variables explain or account for 52 per cent of the variation in the surface aggregation between samples. The significance of each factor within the equation was determined by an F -ratio between the residual error of the equation with all factors present and the residual error of the equation with the given factor omitted (Table 3). As expected, geologic rock type (G) was most important, the level of significance being beyond 0.5 per cent. Vegetative cover type (C) was significant at 5 per cent, and elevation at 10 per cent. Zone was not significant at 5 per cent by itself; however the geology \times zone interaction was significant at 10 per cent.

The partial coefficients given in Table 4 indicate the effect of each variable on the erodibility index. These effects are determined for each variable, with all others held constant. Thus they are true partials and can be directly

3. Partial Regression Coefficients and Significance of Relation of Two Soil Erodibility Indexes—Surface-Aggregation Ratio (*S/A*)—Regression Ratio (*DR*)—and Regression Coefficients for Surface and Aggregated Silt plus Clay and to Geologic Rock Type, Vegetation Zone, and Elevation, Northern California

Partial Regression Coefficients				
Variable	<i>S/A</i>	Surface	Agg. Si + Cl	<i>DR</i>
Rock Type (<i>G</i>)				
	43.1	3.41	-5.53	9.10
	-25.5	-1.25	3.88	0.66
	-35.0	-2.70	2.34	-10.52
Geographic Zone (<i>C</i>)				
Central	-29.7	-2.24	2.19	-2.70
	22.1	-0.21	-7.36	11.62
Coastal	-16.5	-0.90	1.80	-4.80
Sierra	-9.2	0.76	4.06	-4.62
	-37.0	-0.63	0.19	-4.80
Vegetation Cover (<i>E</i>)				
	12.9	0.29	-2.23	0.03
	0.3	-0.10	0.08	-1.19
	-8.1	-0.12	1.33	0.62
Geologic Rock Type (<i>G</i>)				
Coastal	7.7	0.58	1.86	-3.86
Nevada	-1.4	0.13	0.91	1.37
Coastal	-1.8	-0.34	-1.53	0.24
Vegetation Cover (<i>E</i>)				
	26.0	0.11	-7.10	5.41
	-9.9	0.02	3.94	-1.67
	-0.6	-0.22	-0.66	-0.31
	-6.1	0.26	0.55	-1.76
Test of Significance of Independent Variables (<i>F</i> Ratios)				
Variable	<i>S/A</i>			<i>DR</i>
Geologic Rock Type (<i>G</i>)				
	8.77*			9.36*
	2.93†			0.28
	2.24‡			1.90
Geographic Zone (<i>C</i>)				
	1.71‡			1.44
	1.64‡			1.33
	0.12			1.50

* Significant at 0.005.
† Significant at 0.05.
‡ Significant at 0.10.

compared with one another to yield litho-, vegetation, and elevation sequences.

Some generalizations may be made on the relative erodibility of soils formed on the geologic rock types by substituting in the equation the values of Table 4. The coefficients of geology (*G*) in Table 4, if arranged in order of magnitude, give a lithosequence of erodibility. Soil developed on acid igneous rocks has a high erodibility (*S/A* value). If we consider the acid igneous type, averaged for all combinations of vegetation, elevation, and zone (equal areas), we find the *S/A* value equals 131 ($77.38 + 43.1 + 2.4 + 1.7 + 1.5 + 6.7 - 1.5 = 131$). (The last five items 2.4, + 1.7, etc., are small adjustments of the regression constant 77.38, needed because all combinations of vegetation types, elevations, and zones were not sampled with exactly the same frequency.) Similarly, the least erodible geologic type is serpentine, with an *S/A* value of 47. The basic igneous type (basalt) has an *S/A* ratio somewhat higher, 58. These values correspond quite closely to those obtained in Oregon [Anderson, 1954]: 164 for acid igneous and 59 for basalt. Thus both studies indicate that soils developed on acid igneous rocks are about $2\frac{1}{2}$ times as erodible as soils developed on basalt.

Partials for vegetation type from Table 4 give a vegetation sequence of erodibility: grass cover is associated with the least erodible soils, forest cover with soils intermediate in erodibility, and brush cover with the most erodible. Vegetation has not only an over-all effect but also a variable effect on the index, depending upon the type of rock over which the cover is found, as shown by the *G* × *C* interaction values. The highly erodible brushland soils are most erodible when developed on soft sediments and alluvium.

The partial regression coefficients for geographic zone were expected to be associated with other climatic effects; yet zone by itself was not significant at 10 per cent. The importance of zone did show up in the interaction with geology, however, which was significant at 10 per cent. The highly erodible acid igneous type was most erodible in the Central Coastal area; the moderately erodible soft sediments were less erodible in the North Coastal area but highly erodible in the Sierra. With these excep-

tions, erodibility of geologic types is quite consistent between zones.

Other relations to geology. The dispersion ratio DR can be predicted directly from the

geologic rock type (G). The dispersion ratio equals 52.4 plus the soil effect (G) from the column of Table 3; for example, the average dispersion ratio for soils developed on acid

TABLE 4. Regression Equation and Coefficients Relating the Soil Erodibility Index—Surface Aggregation Ratio, S/A —to the Geologic Soil Type, Vegetation Cover, Elevation, and Zone and Their Interactions in Northern California

Prediction equation: $S/A = 77.38 + G + C + Z + E + (G \times Z) + (G \times C)$
 S/A coded by factor of 100.

Geology (G) Signif. beyond 0.005%	Partial Regression Coefficients		
	Vegetation Cover Type (C) Signif. at 5%	($G \times Z$) Signif. at 10%	($G \times C$) Signif. at 10%
$G_1 = +43.1$ Acid	$C_1 = +0.3$ Forest	$G_1 \times Z_1 = +0.2$	$G_1 \times C_1 = -$
$G_2 = -25.5$ Basic	$C_2 = +12.9$ Brush	$G_1 \times Z_2 = -14.1$	$G_1 \times C_2 = -$
$G_3 = -35.0$ Serp.	$C_3 = -8.1$ Grass	$G_1 \times Z_3 = +33.9$	$G_1 \times C_3 = +$
$G_4 = -29.7$ Meta.			
$G_5 = +22.1$ Schist	$Av = +1.7$	$Av = +6.7$	$Av = -$
$G_6 = -16.5$ Hard Sed.		$G_2 \times Z_1 = +11.3$	$G_2 \times C_1 = +$
$G_7 = -9.2$ Soft Sed.		$G_2 \times Z_2 = -10.2$	$G_2 \times C_2 = -$
$G_8 = +37.2$ Allu.		$G_2 \times Z_3 = 0$	$G_2 \times C_3 = -$
$Av = -1.7$			
		$Av = +0.4$	$Av = -$
		$G_3 \times Z_1 = +6.5$	$G_3 \times C_1 = +$
		$G_3 \times Z_2 = +5.1$	$G_3 \times C_2 = -$
		$G_3 \times Z_3 = -21.7$	$G_3 \times C_3 = +$
		$Av = -3.4$	$Av = +$
		$G_4 \times Z_1 = +2.8$	$G_4 \times C_1 = -$
		$G_4 \times Z_2 = -1.5$	$G_4 \times C_2 = +$
		$G_4 \times Z_3 = 0$	$G_4 \times C_3 = +$
		$Av = +0.4$	$Av = -$
		$G_5 \times Z_1 = +8.4$	$G_5 \times C_1 = -2$
		$G_5 \times Z_2 = -11.2$	$G_5 \times C_2 = -2$
		$G_5 \times Z_3 = 0$	$G_5 \times C_3 = +2$
		$Av = -0.9$	$Av = -$
		$G_6 \times Z_1 = -0.8$	$G_6 \times C_1 = +1$
		$G_6 \times Z_2 = 0$	$G_6 \times C_2 = -$
		$G_6 \times Z_3 = +1.7$	$G_6 \times C_3 = -$
		$Av = +0.3$	$Av = +$
		$G_7 \times Z_1 = -28.0$	$G_7 \times C_1 = +1$
		$G_7 \times Z_2 = +44.8$	$G_7 \times C_2 = +3$
		$G_7 \times Z_3 = -9.8$	$G_7 \times C_3 = -1$
		$Av = +2.3$	$Av = +1$
		$G_8 \times Z_1 = +2.0$	$G_8 \times C_1 = -6$
		$G_8 \times Z_2 = +18.6$	$G_8 \times C_2 = +10$
		$G_8 \times Z_3 = -25.5$	$G_8 \times C_3 = -$
		$Av = -1.6$	$Av = +1$
		($G \times Z$) $Av = +0.2$	($G \times C$) $Av = +$

ks is 52.4 plus 9.1, or 61.5. Similarly, the and aggregated silt plus clay can be ed, the constants for these being 10.2 .7, respectively, and the regression co- s coming from the second and third s of Table 3.

tions to combined chemical and soil- y factors. Would estimation of the sur- gregation ratio (S/A) be improved by n of variables of chemical characteristics e of soil formation? In a study by *Wallis evan* [1961] the amount of calcium and ium absorbed on the clays of 20 samples soils was found to be related to the sur- gregation (S/A) ratio.

tested this relationship: S/A as a func- calculated S/A from Table 4 (s/a), the n plus magnesium milliequivalent per ams ($Ca + Mg$), and the calcium plus ium squared. The prediction equation became

$$= 156.4 + 1.06(s/a)$$

$$17.11 (Ca + Mg) + 0.4 (Ca + Mg)^2$$

regression coefficients were all significant; ed variance (R^2) was 63 per cent. The n of the soil-forming factors of geology, tion type, elevation, and zone improved edictability of erodibility reported by and *Stevan* [1961] by 30 per cent. Better tion of erodibility may come from in- ed knowledge of the base status of soil de- id under different conditions of soil-form- ctors. For example, acid igneous rock that are high in the calcium feldspars may ifferent erodibility than the same general ype that is high in sodium feldspars.

CONCLUSIONS AND APPLICATIONS

conclude that both soil erodibility indexes nificantly related to soil-geologic rock and that the surface-aggregation ratio is elated to vegetation and geographic zone rthern California. Such indexes should e useful in evaluating sediment problems ediment control in the state.

The partial regression coefficients can be y compared with one another or used n any combination desired to predict odibility. If important factors are left

out, however, an individual prediction con- siderably in error may result. The prediction equation can be used for evaluating the relative soil erodibility for a watershed by taking a sample of points in the watershed, or it can be used for an entire watershed, by taking the average of each of the conditions present and weighting each factor by its appropriate pro- portions. It should be remembered that the in- dex is for standard slope of 20 to 30 per cent gradient, west exposure, and full vegetative cover; consequently, slope, aspect, and cover condition variables would logically be included in any analysis in which the indexes are used.

2. When the soil erodibility is considered with other factors in erosion—the intensity and frequency of rainfall and land use and condition, for example—then erosion may be predicted and erosion hazard assigned to land areas [*Anderson*, 1957]. The exact relation will depend on the measure of erosion used. *Anderson* [1954] showed that suspended sediment varied with the S/A ratio in such a manner that logarithm suspended sediment varied as $0.482 (S/A)/100$; so soil developed on acid would produce 2.3 times as much erosion (suspended sediment) as soil from basalt rock. For another measure of sedimentation, the number of days per year with turbid streamflow (> 27.5 ppm), the same proportion, $0.48 (S/A)/100$, appeared. We may expect that these erodibility indexes will prove useful in still other sedimentation problems—reservoir deposition and channel scour—and perhaps in channel geometry problems such as those studied by *Schumm* [1960].

Acknowledgments. This study was conducted by the Pacific Southwest Forest and Range Experiment Station, Berkeley, California, with the cooperation of the State Department of Water Resources, Sacramento, California. The cooperation of the University of California in making available laboratory facilities for the analyses is also gratefully acknowledged.

REFERENCES

- Anderson*, H. W., Physical characteristics of soils related to erosion, *J. Soil and Water Conserv.*, 6(3), 129-133, 1951.
- Anderson*, H. W., Suspended sediment discharge as related to streamflow, topography, soil, and land use, *Trans. Am. Geophys. Union*, 35(2), 268-281, 1954.
- Anderson*, H. W., Relating sediment yield to wa-

- tershed variables, *Trans. Am. Geophys. Union*, 38(6), 921-924, 1957.
- André, J. E., Establishment report, soil erodibility samples collected and analyzed, northern California wildlands, Watershed Management Research Division, Pacific Southwest Forest and Range Exp. Sta., Berkeley, Calif., 39 pp. (Processed.) April 4, 1960.
- Baver, L. D., Some soil factors affecting erosion, *Agr. Eng.*, 14, 51-52, 1932.
- Disaker, E. G., and R. E. Yoder, Sheet erosion studies on cecil clay, *Alabama Agr. Expt. Sta. Bull.* 245, 1936.
- Jenny, Hans, Arrangement of soil series and types according to functions of soil forming factors, *Soil Sci.*, 61(51), 375-391, 1946.
- Lutz, J. F., The physio-chemical properties of soils affecting erosion, *Missouri Agr. Expt. Sta. Research Bull.* 212, 1934.
- Middleton, H. E., Properties of soils which influence soil erosion, *U. S. Dept. Agr. Tech. Bull.* 178, 16 pp., Mar. 1930.
- Schumm, S. A., The effect of sediment type on shape and stratification of some modern fluvial deposits, *Am. J. Sci.*, 258, 177-184, 1960.
- Smerdon, E. T., and R. P. Beasley, The tractive force theory applied to stability of open channels in cohesive soils, *Univ. of Missouri Agr. Expt. Sta. Research Bull.* 715, 36 pp., 1959.
- Wallis, James R., and Lee Stevan, Erodibility of some California wildland soils related to their metallic cation exchange capacity, *J. Geophysical Research*, 66, 1225-1230, 1961.
- Yoder, R. E., A direct method of aggregate analysis of soils and a study of the physical nature of erosion losses, *J. Am. Soc. Agron.*, 28, 337-344, 1936.

(Manuscript received May 6, 1961;
revised August 3, 1961.)

Evaluating Wells and Aquifers by Analytical Methods

WILLIAM C. WALTON AND WILLIAM H. WALWER

*Illinois State Water Survey
Urbana, Illinois*

Abstract. The practical application of scientific analytical methods to ground-water management problems in Illinois is described in this report. The actual ground-water flow is simulated by a model aquifer having straight-line boundaries, an effective width, length, and thickness, and sometimes a confining bed with an effective thickness. The hydraulic values of the model aquifer and its confining bed, the image-well theory, and ground-water flow are used to construct a mathematical model which provides a means of evaluating performance of wells and aquifers. Records of past pumping and water levels establish values of the mechanism as a model of the response of an aquifer to heavy pumping. The demonstration of the applicability of model aquifers and mathematical models, the use of ground-water development in the Chicago region in northeastern Illinois and the Taylorville, Illinois, Assumption, and Pekin areas in central Illinois are described. Model aquifers for the study areas range from a semi-infinite permeable strip of sand and gravel 84 miles wide and 1000 feet thick to a semi-infinite permeable strip of sand and gravel 300 feet wide and 3.5 feet thick. Practical estimated yields of the aquifers range from 46 million gallons per day to 16,000 gallons per day.

Introduction

In the last few years there has been increasing recognition of the need for refined data concerning available water resources and their management. One of our aquifers continues to be actively used for the needs of industrial, urban, and suburban expansion. As ground-water development increases, well owners become more dependent on the response of aquifers to heavy pumping. Initially they were concerned with the detection and exploration of the aquifer for available sources. As the response of aquifers to heavy pumping increases, they are becoming more concerned with the management of the aquifer. Before ground-water resources can be managed they must be quantitatively appraised.

Increasing numbers of engineers and geologists are being called upon to estimate how much water is available for development. It will be the joint presence of exploration and water users are continuously asking questions as to how available resources can be managed. Significant changes in the study in ground-water hydrology are being made to permit hydrologists to cope with these questions.

Emphasis is now being placed on the quantitative description of the geologic and hydrologic parameters affecting the water-yielding capacity of wells and aquifers. Scattered and somewhat disconnected studies of wells and parts of aquifers are being put together into regional type studies to facilitate proper development and management of entire ground-water reservoirs. Hydrologists are coming to grips with the complexity of geologic conditions and are devising methods for simulating nature so that wells and aquifers can be evaluated.

Research is being directed toward developing analytical methods and analog models which will take into account existing ground-water flow systems. Analytical methods have heretofore been applied to only relatively uniform aquifers with simple geometry and have been used primarily for problems involving small parts of aquifers or aquifers of small areal extent. Analog models, especially electrical models, may be more realistic and adaptable than analytical methods. However, the realism and versatility of analog models will depend largely upon the availability of abundant and accurate basic data. In some cases basic data are not sufficient to warrant a rigorous description of complex aquifer conditions.

tions and analytical methods may be as useful as analog models.

Recognized departures from ideal conditions do not necessarily dictate that analytical methods be rarely used. Such departures emphasize the need for sound professional judgment in the application of mathematical formulas to existing geologic conditions and in properly qualifying results according to the extent of departures. With appropriate recognition of hydrogeologic controls there are many practical ways of circumventing analytical difficulties posed by complicated field conditions. Many aquifers can be highly idealized with little sacrifice in accuracy of analysis.

MODEL AQUIFERS AND MATHEMATICAL MODELS

In applying analytical methods to field problems the hydrogeologic boundaries of the aquifer evident from areal studies must be idealized to fit comparatively elementary geometric forms such as wedges and infinite or semi-infinite rectilinear strips. Boundaries are assumed to be straight-line demarcations. The gross hydraulic properties of the aquifer and confining bed, if present, are considered in evaluating the effects of boundaries, and the detailed hydraulic property variations are considered in estimating interference between wells.

Actual ground-water conditions are simulated with model aquifers which have straight-line boundaries and an effective width, length, and thickness. The aquifer is sometimes overlain by a confining bed which has an effective thickness.

Mathematical models are based on the hydraulic properties of model aquifers, the image-well theory, and ground-water formulas. Problems associated with hydrogeologic boundaries are simplified to the consideration of an infinite aquifer in which real and image wells operate simultaneously. The effects of real and image wells are computed with appropriate ground-water formulas.

Most hydrogeologic boundaries are not clear-cut straight-line features but are irregular in shape and extent. However, it is generally permissible to treat boundaries as straight-line demarcations because irregularities are often small when compared with the areal extent of most aquifers. It should be recognized that idealized mathematical models describe the

drawdown least accurately in the immediate vicinity of boundaries. The greater the distance to the boundary from the observation point the smaller will be the error involved by approximation.

Records of past pumpage and water level may be used to establish whether assumed mathematical models satisfy the hydrogeologic limits of an aquifer. If computed and actual water-level declines agree, the mathematical model provides a means of evaluating the performance of wells and aquifers.

For a demonstration of the applicability of model aquifers and mathematical models a case history [see *Walker and Walton, 1961*] of ground-water development in east-central Illinois is described below in detail.

INVESTIGATION OF THE ARCOLA AREA, ILLINOIS

The village of Arcola is located in the southern part of Douglas County, 23 miles south of the city of Champaign and 45 miles east-southeast of the city of Decatur. The municipal water supply is obtained from wells in the unconsolidated deposits within and near the city corporate limits.

Based on geologic studies made by the Illinois State Geological Survey, the unconsolidated glacial deposits in the Arcola area are mainly Wisconsinan and Illinoian in age and range in thickness from 80 to 125 feet. As shown in Figure 1 these deposits consist primarily of ice-laid till with some permeable water-laid silt sand, and gravel outwash. The thicker sections of glacial material are contained in a narrow bedrock valley cut in Pennsylvanian bedrock which consists mainly of shale. The thicker and more permeable outwash materials, hereafter referred to as the aquifer, are generally found in the lower part of the drift and are Illinoian in age.

The thick upper unit of the Wisconsinan glacial till, which occurs from the surface to an average depth of 60 feet, contains a high percentage of silt and clay. The lower Wisconsinan unit and the Illinoian deposits immediately overlying the aquifer, hereafter referred to as the confining bed, contain sand lenses with sandy till. The aquifer contains a large amount of fine sand and silt, and its permeability is not great.

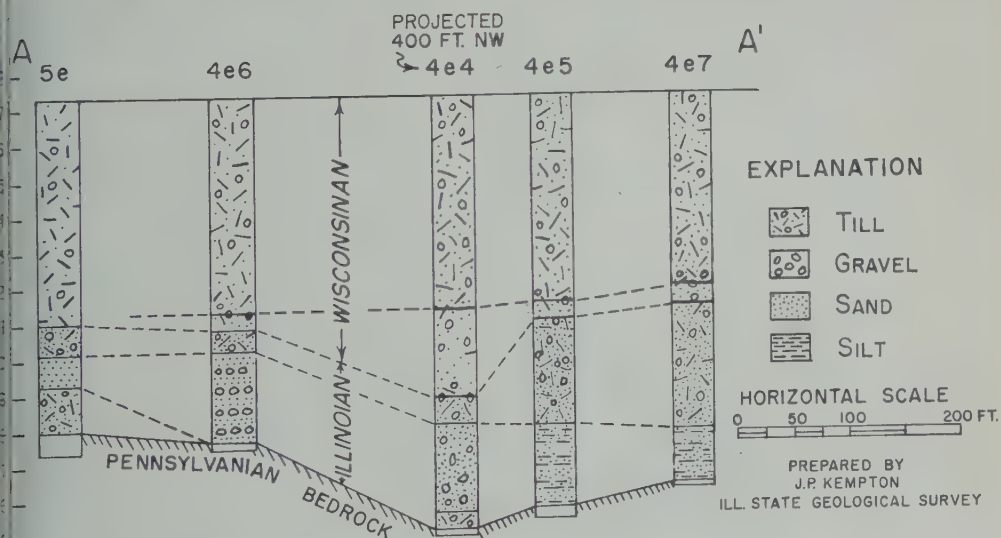
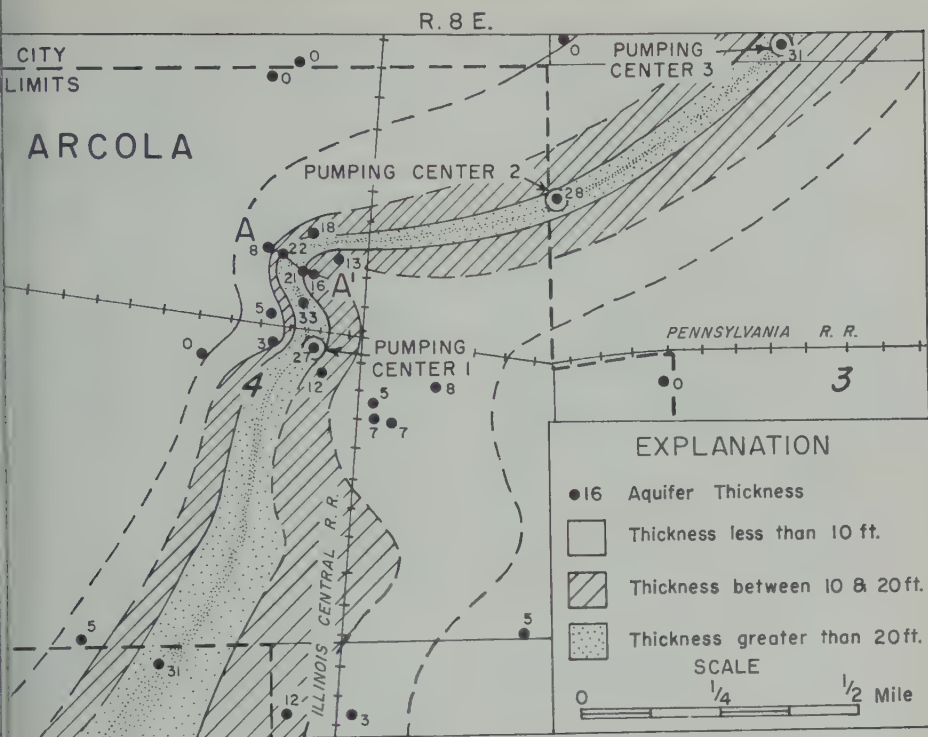


Fig. 1. Map and geologic cross section showing thickness and areal extent of aquifer at Arcola, Illinois.

geologic cross section and aquifer thickness map shown in Figure 1 were drawn from available drillers' logs of wells and test logs. As is often the case, data are not sufficient for a rigorous description of the areal extent of

the aquifer. Analysis of existing geologic information suggests that the aquifer occurs as a thin and narrow strip of permeable sand and gravel exceeding 20 feet in thickness in many places and trending from northeast to southwest

through Arcola. The more permeable part of the aquifer that is suitable for development ranges in width from about 800 to less than 200 feet. Water occurs under leaky artesian conditions, and recharge is derived from the vertical leakage of water through the confining bed into the aquifer.

During the period 1940 to 1955, five pumping tests were made at Arcola to determine the hydraulic properties of the aquifer. Based on pumping-test data the coefficient of permeability P of the aquifer ranges from 280 to 660 gallons per day (gpd) per square foot and the coefficient of transmissibility T ranges from 2200 to 18,000 gpd per foot. The smaller values of T and P reflect thinner and less permeable deposits near the edge of the aquifer; the larger values reflect much thicker and more permeable deposits near the center of the aquifer. The average coefficient of storage of the aquifer is 0.001.

The municipal water supply has been obtained from three pumping centers shown in Figure 1. Average daily ground-water withdrawal increased from 20,000 gallons in 1891 to 146,000 gallons in 1959.

Past records of water-level decline and pumpage and analytical methods were used to determine the practical sustained yield of the aquifer. The practical sustained yield is here defined as the maximum amount of water that can be continuously withdrawn from existing wells without eventually lowering water levels below tops of screens.

An idealized model aquifer that duplicates hydrogeologic conditions in the Arcola area was created. Major factors considered in creating the model aquifer were (1) the external barrier boundaries (bedrock walls) are tapered and irregular in shape, (2) the bedrock walls are not entirely impervious and some subsurface flow will occur from the bedrock into the aquifer, the effective boundaries not being likely to coincide exactly with physical boundaries, (3) the hydraulic properties vary from place to place and are highly variable in the proximity of boundaries but are fairly uniform on a gross basis, and (4) the vertical permeability and thickness of the confining bed vary from place to place but are fairly uniform on a gross basis. With these factors taken into account, the results of geologic and hydrologic studies indicate that it is possible to simulate complex aquifer

conditions with an infinite strip of sand and gravel which is 400 feet wide, 20 feet thick, bounded on the sides and bottom by impermeable material, and overlain by a confining layer 70 feet thick. The model aquifer and its orientation with respect to Arcola are shown in Figure 2a. The average coefficients of transmissibility and storage of the model aquifer are 10,000 gpd/ft and 0.001, respectively.

Most drawdown data collected during pumping tests are affected by barrier boundaries, and it is impossible to isolate the effects of the leakage through the confining bed. Although the vertical permeability of the confining bed cannot be determined from pumping-test data, it can be estimated with the model aquifer.

The water-level decline in an observation well near pumping center 1 was computed using a mathematical model based on the model aquifer, the image-well theory [see *Ferris*, 1959], the steady-state leaky artesian formula described by *Jacob* [1946], estimated pumpage data, and several assumed values of the vertical permeability of the confining bed. The computed decline was then compared with the actual decline and that vertical permeability which gave the computed decline equal to actual observed decline was assigned to the confining bed.

The pumping center, the observation well, and the image wells associated with the boundaries of the model aquifer were drawn to scale on a map. The boundaries are parallel; therefore an image-well system extending to infinity is required [Knowles, 1955]. However, in practice it is only necessary to add pairs of image wells until the effect of the next pair has no measurable influence. A semilog distance-drawdown graph based on assumed hydraulic properties of the model aquifer and its confining bed and past ground-water withdrawals was constructed. The virtual radius r_w of the cone of depression, the distance from the pumped well beyond which drawdown is not measurable, was determined from the graph, and image wells at greater distances than r_w were not considered. The map showing the location of the pumping center, observation well, and image wells and the distance-drawdown graph as shown in Figure 2b constitute the mathematical model.

The distances between the observation well, the pumping center, and the image wells were

from the map. The computation of the level decline in the observation well was on past pumping data; the distance-drawdown graph was used to compute the decline of the real and image wells.

The observed decline in the observation well at a pumping rate of 115 gpd is 42 feet. A level decline of 42 feet was computed with a distance-drawdown graph based on a vertical permeability of 0.04 gpd/ft². Therefore, a vertical permeability of 0.04 gpd/ft² was assigned to the confining bed overlying the model aquifer. To test the mathematical model, the drawdowns recorded in pumping centers 2 and 3 and by pumping center 1 were compared with drawdowns at the pumping centers computed with the mathematical model. Actual drawdowns in pumping centers 2 and 3 of 30 and 19 feet, respectively, are within a few per cent of the computed declines of 32 and 19 feet. This close agreement between computed and actual decline indicates that the model aquifer and the mathematical model closely describe the geologic conditions at Arcola. It is reasonable to assume that the model aquifer and the mathematical model may be used to predict with reasonable accuracy the effects of future ground-water development and the practical sustained yield of the aquifer.

The mathematical model is based on a particular combination of aquifer boundaries and hydraulic properties. There are probably other mathematical models involving several slightly different combinations of parameters which would duplicate aquifer conditions.

In 1959 the multiple-well system at Arcola consisted of three wells ranging in depth from 100 to 122 feet and spaced 2500 feet apart in the more permeable parts of the aquifer as shown in Figure 2a. Available drawdowns in the observation wells, assuming pumping levels above the water table, range from 66 to 75 feet. Computations with the mathematical model indicate that the practical sustained yield of the existing 3-well system is about 137 gallons per minute or 200,000 gpd. The practical sustained yield can be developed by pumping production wells 1 and 3 at 57 gpm and production well 2 at 23 gpm. Variations in hydraulic properties were considered in computations of drawdowns by pumping the production wells themselves. The practical rate of ground-water withdrawal in-

creased from 115,000 gpd in 1957 to 146,000 gpd in 1959. If pumpage continues to increase at this rate in the future, it is estimated that the practical sustained yield of the existing 3-well system will be exceeded by 1963.

The Arcola case history is suggestive of how analytical methods can be utilized to evaluate wells and aquifers so that available ground-water resources can be properly managed. By checking computed performance of wells and aquifers with records of past pumpage and water levels, the hydrologist is assured of reasonably accurate solutions. The time-consuming computations associated with mathematical models can in most instances be reduced to simple standardized procedures which can be inexpensively handled by means of digital computing machines [Walton and Neill, 1961].

EVALUATION OF SEVERAL AQUIFERS IN ILLINOIS

To date, the principles outlined above have been applied to aquifer conditions in five areas in Illinois in addition to the Arcola area. The practical sustained yields of well fields and aquifers in the Chicago region in northeastern Illinois and in the Taylorville, Tallula, Assumption, and Pekin areas in central Illinois have been evaluated. Model aquifers and mathematical models for these areas are shown in Figures 2, 3, and 4.

Chicago region. The Cambrian-Ordovician aquifer [Suter, Bergstrom, Smith, Emrich, Walton, and Larson, 1959] is the most highly developed aquifer for large ground-water supplies in the Chicago region. The Cambrian-Ordovician aquifer is encountered at an average depth of about 500 feet below the land surface at Chicago; it has an average thickness of 1000 feet and is composed chiefly of sandstones and dolomites. The Maquoketa formation consisting largely of shale overlies the Cambrian-Ordovician aquifer and confines the water in the deep aquifer under leaky artesian conditions. The Cambrian-Ordovician aquifer receives water from overlying glacial deposits in areas averaging 47 miles west of Chicago where the Maquoketa formation is absent.

Based on the results of 63 pumping tests and other studies the coefficients of transmissibility and storage of the Cambrian-Ordovician aquifer and the vertical permeability of the Maquoketa formation are fairly uniform throughout

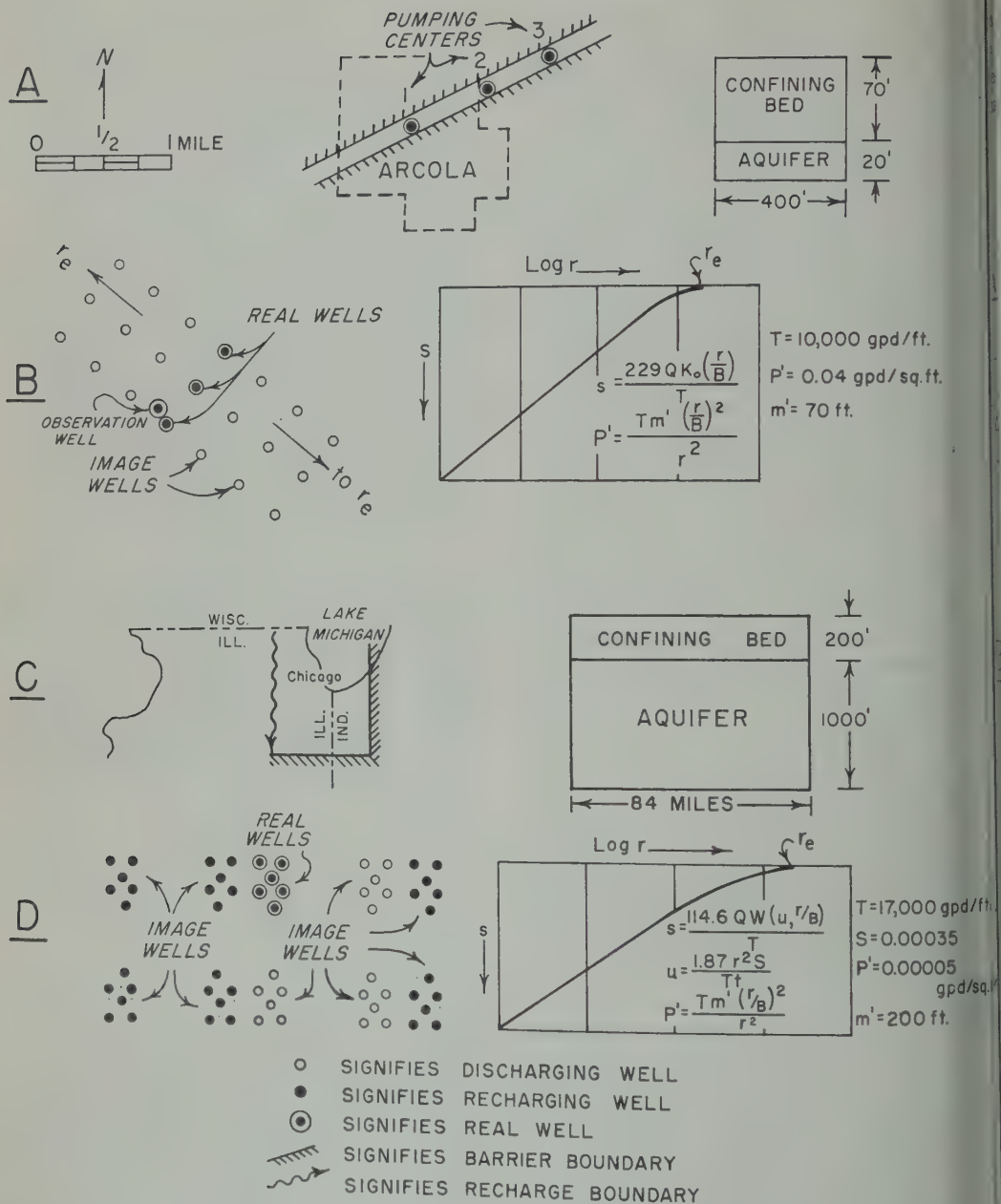


Fig. 2. Model aquifers and mathematical models for Arcola area (A and B) and Chicago region (C and D), Illinois.

large areas in northeastern Illinois and average 17,000 gpd/ft and 0.00035 and 0.00005 gpd/ft², respectively [Suter and others, 1959; Walton, 1960]. The coefficient of transmissibility decreases rapidly south and east of Chicago.

The results of geologic and hydrologic studies

indicate that it is possible to simulate the Cambrian-Ordovician aquifer with an idealized model aquifer as shown in Figure 2c. The model aquifer is a semi-infinite rectilinear strip of sandstones and dolomites 84 miles wide and 1000 feet thick. The model aquifer is bounded

recharge boundary 47 miles west of Chi-
and by two intersecting barrier boun-
37 miles east and 60 miles south of
o and is overlain by a confining bed con-

sisting mostly of shale averaging 200 feet thick.
The mathematical model for the Cambrian-
Ordovician aquifer is shown in Figure 2d.

Pumpage of ground-water from the Cambrian-

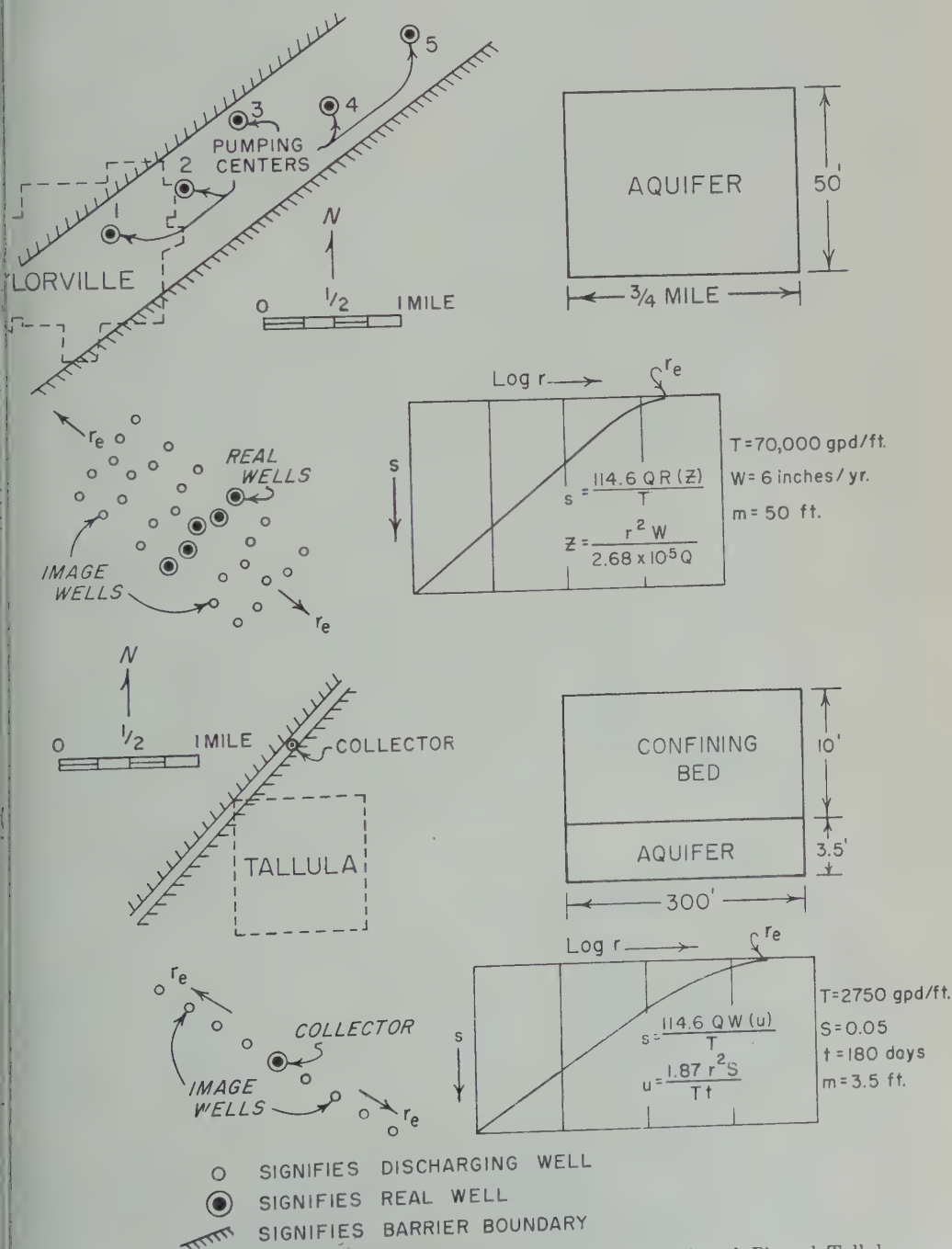


Fig. 3. Model aquifers and mathematical models for Taylorville (A and B) and Tallula (C and D) areas, Illinois.

Ordovician aquifer increased gradually from 200,000 gpd in 1864 to 50 million gpd in 1959. Pumpage is concentrated in six centers as shown in Figure 2d: the Chicago, Joliet, Elmhurst, Des Plaines, Aurora, and Elgin areas. As a result of

heavy pumping, artesian pressure in deep wells declined more than 600 feet at Chicago between 1864 and 1959.

Studies made with the mathematical models show that the practical sustained yield of the

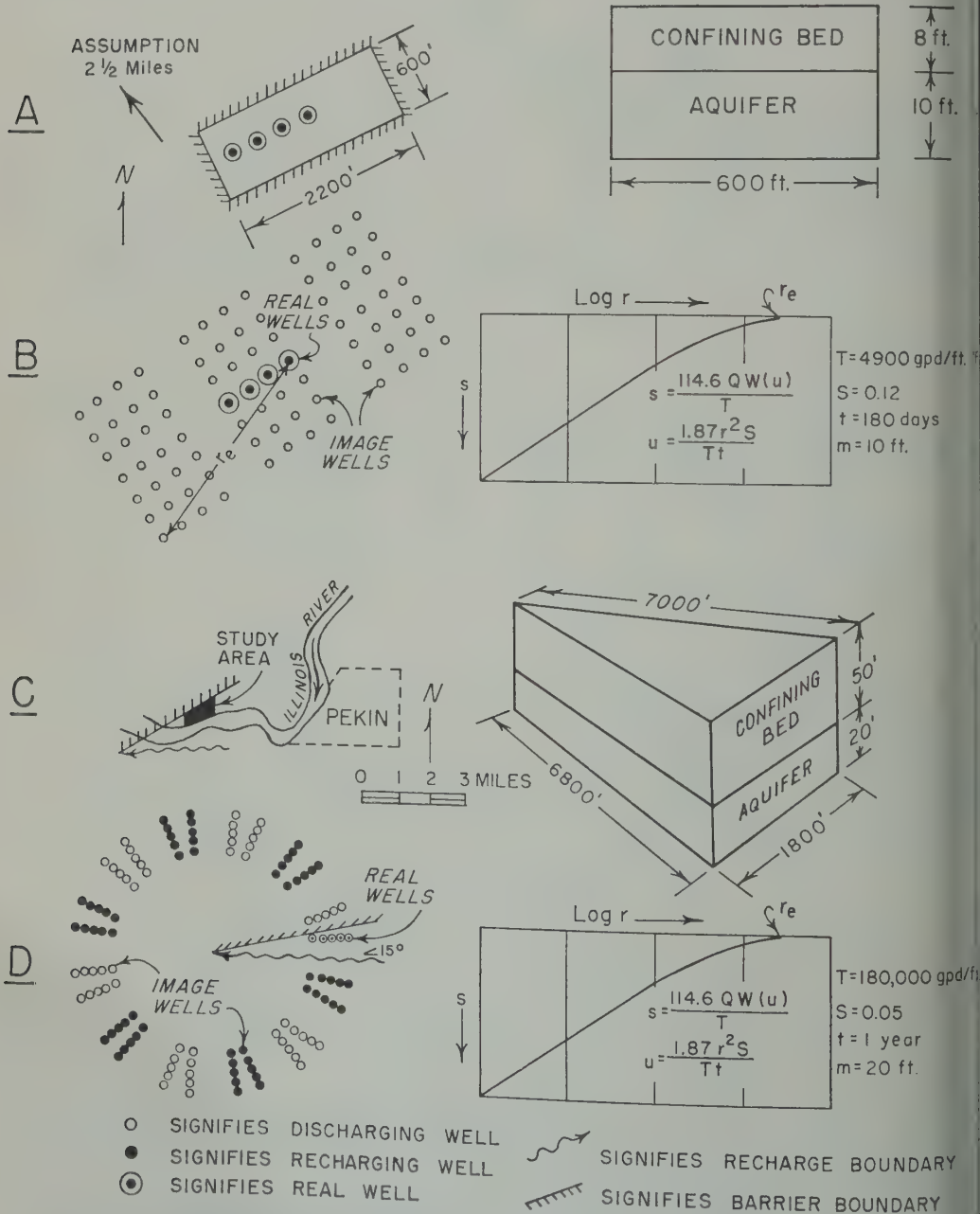


Fig. 4. Model aquifers and mathematical models for Assumption (A and B) and Pekin (C and D) areas, Illinois.

ian-Ordovician aquifer is about 46 miles and is largely limited by the rate at which water can move eastward through the aquifer from recharge areas. The practical sustained yield of the aquifer is here defined as the maximum amount of water that can be consistently withdrawn with the present distribution of pumping centers without eventually causing the most productive and basal water-bearing formation of the Cambrian-Ordovician aquifer. The practical sustained yield of the aquifer was exceeded in 1959, and in a sense all water users in the Chicago region started to borrow water and to borrow water from future generations.

Declines in nonpumping water levels that were expected between 1958 and 1980 at pumping centers were computed by using the mathematical model and assuming that the distribution of pumping remains the same as it was in 1958. Computed declines ranged from 300 feet in the Chicago area to 190 feet at Elgin and averaged 250 feet.

Taylorville area. Municipal and industrial water supplies at Taylorville are obtained from glacial deposits, chiefly Illinoian in age, with a range in thickness from 50 to 180 feet [see *Walton and Walton, 1961*]. The glacial deposits consist of a buried valley cut into relatively impermeable bedrock of Pennsylvanian age and overlain by a complex of ice-laid till, water-laid silt, sand and gravel outwash, and wind-deposited loess and fine sand (loess). The bedrock valley, trending northeast to southwest through Taylorville, was at one time in the past largely filled with glacial till. The outwash sand and gravel aquifer which yields water in large quantities to wells occurs as a fill in a narrow valley within the upper part of the till deposits. The thickness of the deposits range in width from $\frac{1}{2}$ to 1 mile and range in thickness from less than a few feet to 113 feet.

Water occurs in the aquifer under water-table conditions, and the source of recharge is precipitation. The recharge area of the aquifer is bounded approximately by ground-water divides and the edges of the aquifer. It is estimated that an average of 6 inches or 17 percent of the mean annual precipitation reaches the water table in the 6.3-square-mile recharge area.

The hydraulic properties of the aquifer are

known from the results of seven pumping tests. Analysis of data indicates that the coefficient of transmissibility ranges from 34,000 to 130,000 gpd/ft and the coefficient of permeability ranges from 600 to 2200 gpd/ft². The smaller values reflect thinner and less permeable deposits near the edge of the aquifer, whereas the larger values reflect thicker and more permeable deposits near the center of the aquifer. The coefficient of storage of the aquifer averages 0.15.

Total ground-water withdrawals from municipal and industrial wells at Taylorville increased progressively from about 28,000 gpd in 1890 to a maximum of about 3,000,000 gpd in 1953. Pumpage decreased rapidly from 3,000,000 gpd in 1953 to 1,750,000 gpd in 1957. In 1959 total ground-water withdrawal was 1,770,000 gpd. Heavy pumpage, concentrated in five well fields as shown in Figure 3a, caused water levels to decline about 40 feet between 1888 and 1956. Many of the production wells are located in the thinner and less permeable parts of the aquifer.

The idealized model aquifer for the Taylorville area shown in Figure 3a is an infinite rectangular strip of sand and gravel $\frac{3}{4}$ mile wide and 50 feet thick which is bounded on the sides and bottom by impermeable material. The average coefficients of transmissibility and storage of the model aquifer are 70,000 gpd/ft and 0.20, respectively. The mathematical model for the model aquifer is shown in Figure 3b.

Studies made with the mathematical model indicate that, with the present distribution of pumpage and available drawdowns ranging from 40 to 80 feet, the practical sustained yield of the aquifer is 745 gpm or 1,070,000 gpd. About 970 gpm or 1,400,000 gpd can be obtained without excessive drawdown from four wells screened in the thicker and more permeable sections of the aquifer within a 3-mile radius of Taylorville.

Pumpage in 1959 exceeded the practical sustained yield of the aquifer. Computed future water-level declines between 1959 and 1965 indicated that by 1961 pumping levels in many production wells will recede to positions below tops of screens and that in 1965 pumping levels will decline to critical stages several feet below tops of screens.

Tallula area. The municipal water supply for the village of Tallula is obtained from a horizontal collector on the flood plain of a small

creek [see *Walker and Walton, 1961*]. The collector penetrates a thin sand and gravel aquifer that ranges in thickness from 2.5 to 4.5 feet and is encountered at an average depth of 16 feet below land surface. The aquifer is not very permeable, consisting of stratified beds of sand, gravel, and silt in various mixtures. A confining bed averaging 10 feet in thickness and consisting of alluvial clay, silt, and fine sand overlies the aquifer. The sand and gravel aquifer is inferred to be from 150 to 370 feet wide and is contained in a narrow valley cut into relatively impermeable bedrock of Pennsylvanian age.

Recharge is derived chiefly by the vertical leakage of water through the confining bed. Because of the small area and silted condition of the bed of the creek, low stream flow, and the presence of silty materials beneath the stream bed, very little recharge occurs by the induced infiltration of surface water, especially during summer, fall, and winter months. Large amounts of water enter the aquifer through a recharge well connected to a lagoon and located 60 feet from the end of one of the laterals in the collector.

The coefficients of transmissibility and permeability of the aquifer determined from the results of two pumping tests are 2750 gpd/ft and 790 gpd/ft², respectively. Under natural conditions, leaky artesian conditions exist; however, under heavy pumping conditions and during prolonged dry periods the confining bed is partially drained.

Pumpage from the horizontal collector increased from 9000 gpd in 1955 to 29,000 gpd in 1959. The horizontal collector consists of a 6-foot-diameter concrete caisson 26 feet deep from which two horizontal 8-inch-diameter vitrified perforated clay pipe laterals are projected into the aquifer. One horizontal lateral (upper lateral) projects from the caisson at a depth of 18 feet and is 478 feet long. The other horizontal lateral (lower lateral) projects from the caisson at a depth of 21 feet and is 310 feet long. As the result of pumping at a rate of 37,000 gpd during the summer months of 1959, the water levels declined below the top of the upper lateral in the caisson.

The idealized model aquifer for the Tallula area as shown in Figure 3c is a semi-infinite rectilinear strip of sand and gravel 300 feet wide and 3.5 feet thick. It is bounded on the sides and

bottom by impermeable material and is overlain by a confining bed 10 feet thick. The mathematical model for the model aquifer is shown in Figure 3d.

Pumping-test data and the mathematical model were used to compute the radius of a vertical well that would have the same specific capacity as the horizontal collector had. Computations made simulating the collector with a vertical well having a radius of 66 feet and using the mathematical model indicate that the practical sustained yield of the collector is 11 gpm or 16,000 gpd during extended dry periods and 20 gpm or 25,000 gpd during years of normal precipitation.

Assumption area. A new well field was recently developed 2½ miles southeast of the city of Assumption to supplement the municipal water supply. Geologic studies suggest that the aquifer underlying the well field is mainly poorly sorted sand, ranging in thickness from 5 to 15 feet. The aquifer is of limited size; rectangular dimensions are 600 feet wide by 2200 feet long. Clayey materials (confining bed) with an average saturated thickness of 8 feet overlie the aquifer. Water occurs under leaky artesian conditions, and recharge is received chiefly from precipitation by the vertical leakage of water through the confining bed.

Based on the results of two pumping tests, the coefficients of transmissibility and permeability of the aquifer are 4900 gpd/ft and 410 gpd/ft², respectively. The vertical permeability of the confining bed is 0.19 gpd/ft².

The idealized model aquifer for the Assumption area, shown in Figure 4a, is a box of sand 600 feet wide, 2200 feet long, and 10 feet thick. The model aquifer is bounded on the sides and bottom by impermeable material and is overlain by a confining bed with an average saturated thickness of 8 feet. The mathematical model for the model aquifer is shown in Figure 4b.

The response of the aquifer to long-term pumping was studied by means of the mathematical model. Computations based on an average available drawdown of 12 feet indicate that the practical sustained yield of a 4-well system, consisting of wells 6 inches in diameter, 24 feet deep, with 5 feet of screen, and spaced 300 feet apart, is 38 gpm or 55,000 gpd. Gravity drainage of the confining bed and part of the aquifer

extended dry periods was taken into account in estimating the yield of the aquifer. In the study area. Recently an intensive hydrogeologic study was made to determine the feasibility of developing for industrial use a large ground-water supply from unconsolidated deposits in an area along the Illinois River, about 10 miles southwest of the city of Pekin. The study included a test drilling program and a controlled pumping test. The data thus obtained were compared with information from other sources to make an evaluation of the practical sustained yield of the aquifer in the study area and to design a multiple-well system capable of meeting the demand of the industry.

The unconsolidated deposits in the study area consist of recent silty alluvial materials and older outwash of Wisconsinan age. These deposits are contained in a buried valley cut into the underlying bedrock of Pennsylvanian age. The permeable outwash forming the aquifer has a thickness from less than 5 feet to more than 33 feet and consists of stratified beds of silt and sand. Logs of wells and test holes indicate that the boundary marking the limits of the aquifer trends northeast to southwest through the study area. In the proposed well area the aquifer averages 20 feet in thickness and is overlain with fine-grained alluvial materials having an average saturated thickness of 10 feet. The Illinois River, which trends east to west through the study area, has been incised into the alluvial materials but not into the bedrock.

Computations made with pumping-test data indicate that the coefficients of transmissibility and permeability of the aquifer are 180,000 ft and 9000 gpd/ft². During the pumping test, leaky artesian conditions occurred a short time after pumping started, and gravity drainage of the alluvial materials was appreciable during the latter part of the test period, especially in the immediate vicinity of the pumped well. Test data were adversely affected by a boundary (the edge of the aquifer); however, the effects of recharge from the Illinois River caused water levels to stabilize rapidly, indicating a fair connection between the aquifer and the river.

A mathematical model aquifer which simulates the actual hydrogeologic conditions present in the study area is shown in Figure 4c. The mathematical

model for the model aquifer is shown in Figure 4d. Computations made with the mathematical model indicate that 4000 gpm or 5,760,000 gpd can be obtained with maximum drawdown above the top of the aquifer from five wells spaced about 325 feet apart.

CONCLUSIONS

It is often possible to evaluate wells and aquifers with analytical methods by devising approximate methods of analysis based on idealized models of aquifer situations. Case histories of ground-water development indicate that aquifer behavior actually does coincide rather closely with what may be predicted theoretically with mathematical models and mathematical models.

It is recognized that methods of analysis described in this report provide only approximate answers. To quantitatively describe in detail hydrogeologic systems having highly complex geometry and great variations in hydraulic properties, the hydrologist may have to turn to electrical analog models which are more versatile in simulating aquifer conditions. Results obtained from analytical methods and analog models need to be compared to determine the limitations of analytical methods.

It is apparent that quantitative answers depend primarily upon the accurate description of geologic and hydrologic controls. In the future, as the techniques of ground-water resource evaluation are refined, a need for more precise, quantitative data concerning requisite geologic information will develop.

In evaluating aquifers the hydrologist is forced to take a good hard look at the forest as well as the trees and in doing so his understanding of the major factors governing the yields of wells and aquifers is greatly enriched. Too often in the past we have overemphasized the complexity of conditions and have spent most of our time with details, at the expense of understanding the over-all problem. Careful study often greatly reduces the importance of complexities and indicates that stress should be placed on the more dominant factors.

The six case histories described in this report represent a good start in cataloging the practical sustained yields of wells and aquifers in Illinois. With the aid of a digital computer, hypothetical aquifers having external bound-

daries and hydraulic properties that are different from those of the aquifers discussed in the report are being evaluated. The results of continuing studies of actual ground-water development and of the response of hypothetical aquifers to heavy pumping will in the future greatly assist the Illinois State Water Survey in making rough quantitative appraisals of undeveloped aquifers.

REFERENCES

- Ferris, J. G., *Hydrology*, chap. 7, Ground water, edited by C. O. Wisler and E. F. Brater, John Wiley & Sons, New York, 1959.
- Jacob, C. E., Drawdown test to determine effective radius of an artesian well, *Proc. Am. Soc. Civil Engrs.*, 72 (5), 629-646, 1946.
- Knowles, D. B., Ground-water hydraulics (unpublished), Open-file report, U. S. Geological Survey, Washington, 1955.
- Suter, Max, R. E. Bergstrom, H. F. Smith, G. E. Emrich, W. C. Walton, and T. E. Larson, Preliminary report on ground-water resources of the Chicago region, Illinois, *Cooperative Groundwater Rept. 1*, Illinois State Water Survey, 1958.
- Walker, W. H., and W. C. Walton, Ground-water development in three areas of central Illinois, *Rept. of Invest. 41*, Illinois State Water Survey, in press, 1961.
- Walton, W. C., and J. C. Neill, Analyzing ground-water problems with mathematical models and a digital computer, *Assoc. intern. d'hydrol. sci. Publ. 52*, 1961.
- Walton, W. C., Leaky artesian aquifer conditions in Illinois, *Rept. of Invest. 39*, Illinois State Water Survey, 1960.

(Manuscript received April 25, 1961.)

New Evidence for the Impact Origin of the Ries Basin, Bavaria, Germany¹

E. M. SHOEMAKER AND E. C. T. CHAO

*U. S. Geological Survey
Menlo Park, California, and Washington, D. C.*

Abstract. The Ries basin is a shallow, nearly circular depression about 17 miles in diameter that lies between the Swabian and Franconian plateaus of southern Germany. Great masses of breccia and a system of thrust sheets associated with the Ries have been studied by German geologists for about a century. E. Werner and Otto Stutzer suggested that the Ries was an impact crater, but the consensus of the principal investigators has been that it was formed by some sort of volcanic explosion.

The only direct evidence of magmatic activity at the Ries is the presence of glass in scattered patches of a breccia called suevite. Some of the glass has long been recognized as sintered fragments of crystalline rocks. We have found that glasses of various composition coexist in single specimens of suevite. In addition, coesite, a high-pressure polymorph of SiO_2 , and lechatelierite, SiO_2 glass, occur in the sintered rocks in the suevite. The presence of the same phases in sintered rock fragments at Meteor Crater, Arizona, and the coexistence of glasses of different composition suggest that the glassy components of suevite are of impact rather than volcanic origin.

Author's Note: This paper was presented at the Geophysical Laboratory-Lawrence Radiation Laboratory Cratering Symposium held at the Geological Laboratory in Washington, D. C., on August 28 and 29, 1961. The organizing group included Dr. Philip H. Abelson of the Geophysical Laboratory, Carnegie Institution of Washington; Dr. Gerald W. Johnson and Dr. Wilmot N. Hess of the Lawrence Radiation Laboratory, University of California; and Dr. Gordon J. F. MacDonald of the National Aeronautics and Space Administration. Most of the papers presented on fossil meteorite craters, on craters produced by chemical and nuclear explosions, and on lunar craters. The group of papers printed here represents about half of the symposium. The Lawrence Radiation Laboratory plans to publish a complete report of the proceedings of the symposium.

The Ries basin, or Rieskessel, of Bavaria, Germany, is a dissected, nearly circular crater about 18 miles across, about 50 miles east of Ulm (Fig. 1). The town of Nördlingen lies on the rim of the basin. Since its formation, this basin has undergone extensive morphological modification by sedimentation and erosion. Its outer rim now stands about 600 feet higher than the basin floor, and the rim is about 100 feet above the surrounding surface. The

This report concerns work done on behalf of the National Aeronautics and Space Administration, as authorized by the Director, U. S. Geological Survey.

crater and associated structural textures have attracted the attention of German geologists for more than a century [Cotta, 1834; Oberdorfer, 1905; Dorn, 1948], and as long ago as 1904 the crater was suggested to be of impact origin [Werner, 1904]. Review of the published data led Shoemaker to examine the structural evidence in the field in the summer of 1960. Of special interest was the occurrence of a tufflike rock, referred to by the German authors as 'suevit,' which we suspected to be of impact origin.

The discovery of coesite, the high-pressure polymorph of silica from Meteor Crater, Arizona [Chao, Shoemaker, and Madsen, 1960], had given us a new lead toward the recognition of impact craters. It was natural that suevite from the Ries be re-examined. This paper is a preliminary report of our findings.

The Ries basin lies between the Swabian and the Franconian Alb, limestone plateaus underlain by gently dipping and flat-lying Jurassic limestone and Triassic clastic sedimentary rocks. The original topography within the basin comprised a central depression about 7 miles across, now filled with sediments, surrounded by an irregular shelf from which rose scattered hills. Miocene lake beds as much as 1000 feet thick [Reich and Horrix, 1955] and minor deposits of Pleistocene and Recent sediments have been deposited on the crater floor. The original relief

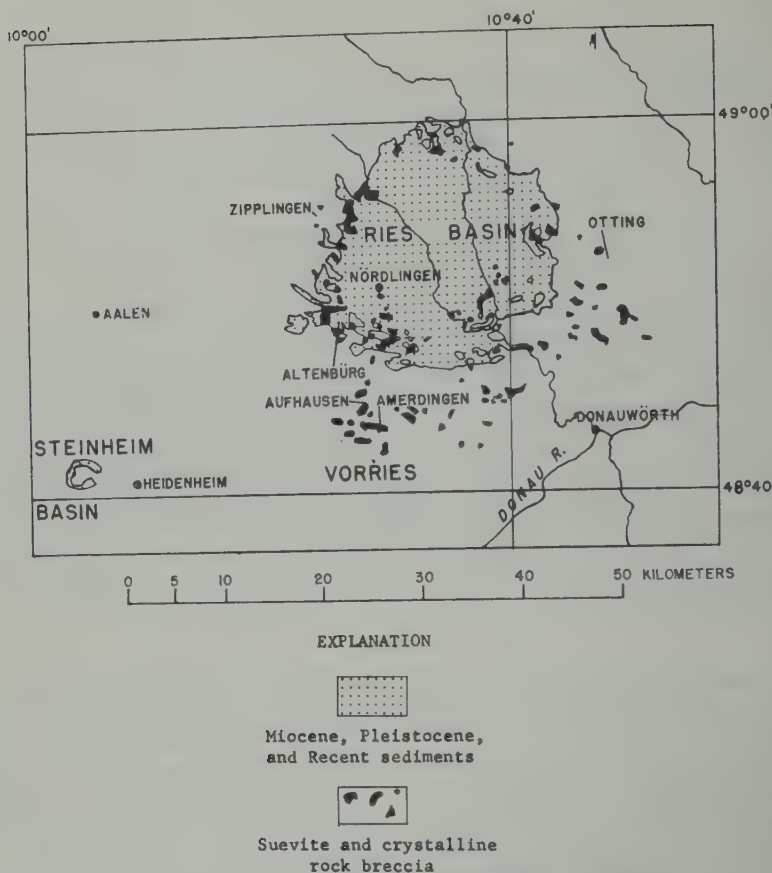


Fig. 1. Index map of the Ries basin, Bavaria, Germany.

between the center of the crater and the crater rim must have been more than 1600 feet. Lake level at some time in the late Miocene was above the present floor of the basin, and extensive deposits of algal tuffas, gastropod coquinas, and other forms of fresh-water limestone occur on scattered hills in the basin and on the lower slopes of the crater walls.

Breccias, chiefly of old crystalline rocks, are exposed on some of the hills that rise above the floor of the Ries today. A variety of granites, interlayered granite and gneiss, aplite, amphibolite, plagioclase amphibolite, greenstone, and schist are the principal crystalline materials. Locally, fragments of these rock types are mixed in varying proportion with fragments of black and red shale of Triassic age and of limestone of Jurassic age. Elsewhere great brecciated masses of Jurassic limestone occur on the floor of the crater.

The walls of the crater and crest of the rim (the 'Schollen- und Schuppen-Zone' of *Bentz* [1927]) are underlain by breccia and an imbricate series of thrust sheets composed chiefly of Upper Jurassic limestones that cap the Alb [*Nathan*, 1925, 1935; *Dehm*, 1932; *Gerstlauer*, 1940; *Schröder and Dehm*, 1950; and *Treibs*, 1950]. Extending tens of miles to the south of the rim, in the region described by *Branco* [1902] as the 'Vorries,' are great masses of limestone breccia resting on the undisturbed limestone cap of the Alb and, at the outer limit, on Oligocene sediments. One far-flung fragment of limestone, measuring nearly half a meter in diameter, was found 40 miles from the crater [*Reuter*, 1925]. In places in the Vorries and the Schollen- und Schuppen-Zone, breccias composed chiefly of Lower Jurassic and Triassic sedimentary rocks or of rocks from the crystalline basement complex rest on the breccias of Upper Jurassic lime-

and the thrust slices. Locally, fragments of basic rocks and crystalline rocks are mixed in equal proportion. Similar breccias were found in a deep drill hole under the lake beds in the center of the Ries. *Branco* [1902] and *Bentz* [1925] long ago concluded that the problem of solving the origin of the Ries lay in determining the mode of emplacement of these breccias.

The most remarkable aspect of the geology of the Ries is the breccia, previously supposed to be of volcanic origin, that has been called *Sauer*, [1901]. The suevite occurs inside the crater, on the rim, and in the Vorries (Fig. 1). It consists of a wide variety of rock fragments, including crystalline rocks from the basement complex, in general shattered or partly sintered, as bombs and smaller fragments of glass. All German authors have agreed are of a fragmental igneous or magmatic origin. The bombs invariably carry fragments of shattered or partly sintered rocks that are recognized as derived from the crystalline basement. Some of the bombs have extraordinary shapes of volcanic ejecta, such as thin sheets that have

been tightly folded or curled on the edges. Although suevite had been studied by *Gümbel* [1870], *Oberdorfer* [1905], and other workers such as *Bentz* [1925], to our knowledge no modern mineralogic or petrographic study had been published before this investigation. Numerous quarries and outcrops of suevite were visited to obtain samples, to study the variations in the character of the breccia, and to learn as much as possible about its structural relations. The first specimen, collected from Otting, and mailed to Chao, proved to contain coesite.

In a quarry at Otting, outside the rim of the crater, suevite is exceptionally well exposed. It forms a patch about 2000 feet across and rests stratigraphically on a breccia composed largely of fragments of red and green claystone set in a dominantly sandy matrix. These pieces of sedimentary rock are derived mainly from Triassic beds, and the material underlying the suevite is a facies of breccia that has been mapped elsewhere as Bunter breccia. The suevite itself forms a nearly massive layer about 30 feet thick with rude stratification or parting parallel to the basal contact. It is composed of a wide

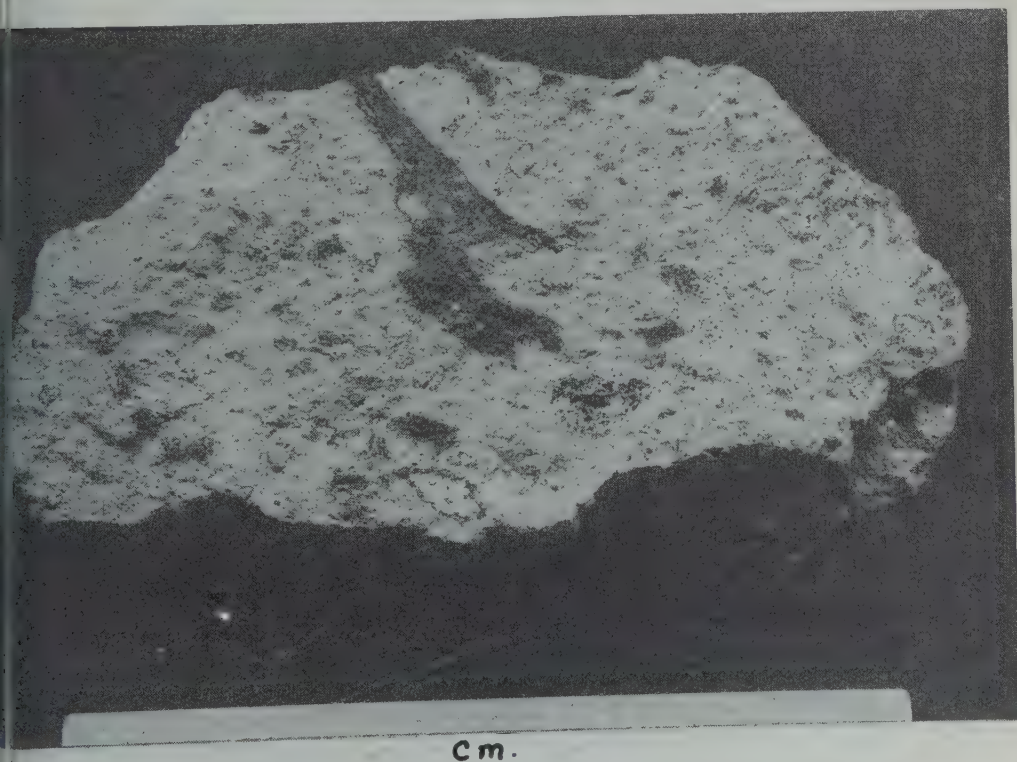


Fig. 2. Photograph of a block of suevite from Otting.

variety of clastic material and bomb- and lapilli-like masses of glass and partly sintered crystalline rocks that range in size from microscopic particles to objects as much as a foot across. Light-colored, partly sintered, fine-grained granitoid and gneissic rocks are the predominant coarser fragments. Under the hand lens the quartz and feldspar grains in most of these fragments are seen to be crushed, and many show varying degrees of sintering and inflation with vesicles. Tightly folded sheets of dark-colored glass (Fig. 2) are common.

One of the light-colored sintered rock fragments was selected for detailed study. It contains more than 80 per cent glass of different kinds, the remainder consisting of a secondary montmorillonitic clay mineral, magnetite, and a small but readily noticeable amount of coesite. Coesite, with high relief, occurs in clear silica glass or lechatelierite (Fig. 3). The silica glass is non-vesicular and may have originated by shock without actual melting. It has an index of 1.462 determined by the immersion method using sodium light at 25°C. These grains of clear silica

glass were separated and examined by X ray. The X-ray diffraction photograph shows only a weak pattern of coesite, a trace of quartz, and the darkened background of the glass. Hand-picked, fractured, partially vitreous, and powdered quartz gave an X-ray pattern of coesite and quartz plus some glass. Several grams of the sintered fragment studied were then dissolved in a weak HF solution to concentrate the coesite. After concentration the coesite was identified both optically by immersion oil and by X-ray diffraction. The predominant glass present in the sample is a clear vesicular glass with an index of refraction of 1.505.

Opaque material in the light-colored fragment (Fig. 3) consists of magnetite. The rounded shape of the opaque material suggests that the pre-existing iron mineral has been fused. It is strongly magnetic, and its powder diffraction pattern is identical to that of magnetite. Spectrographic analysis shows that its nickel content is less than 0.01 per cent.

One of the dark masses of glass studied from the suevite from Otting appears dirty grayish

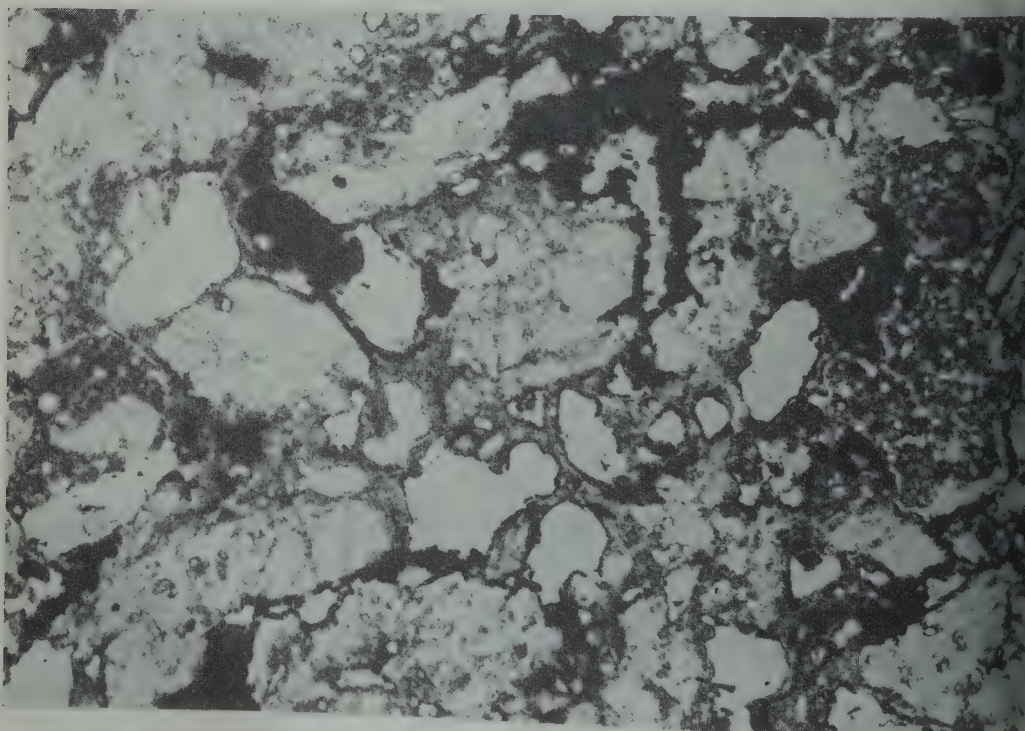


Fig. 3. Photomicrograph of a thin section of suevite from Otting (plane light).

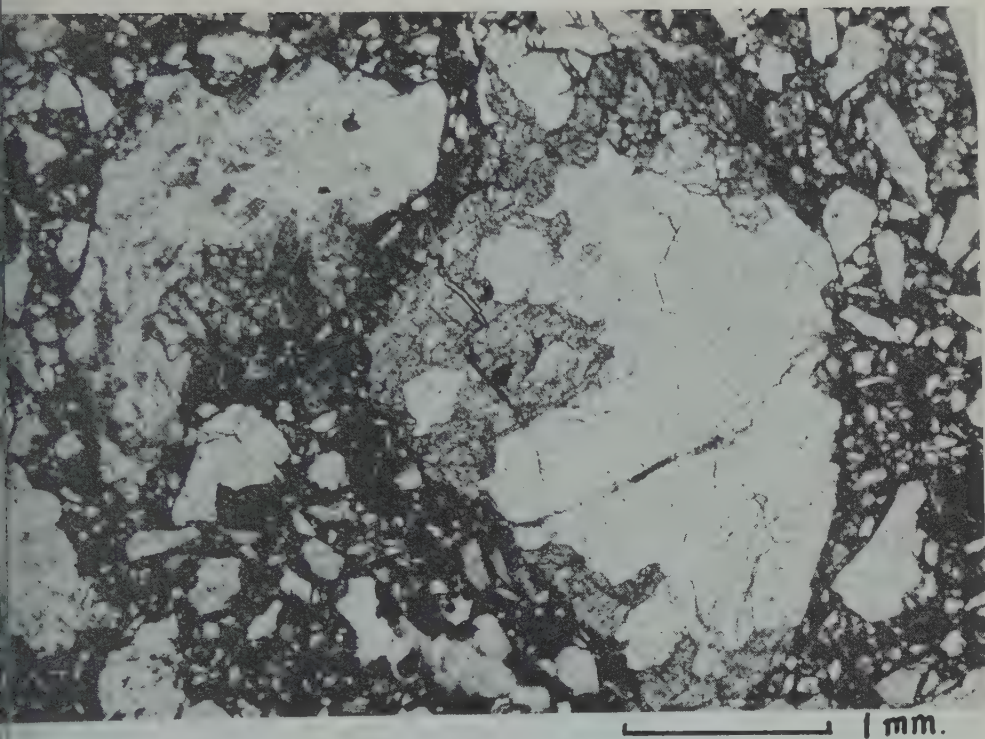


Fig. 4. Photomicrograph of a thin section of suevite from Zipplingen (plane light).

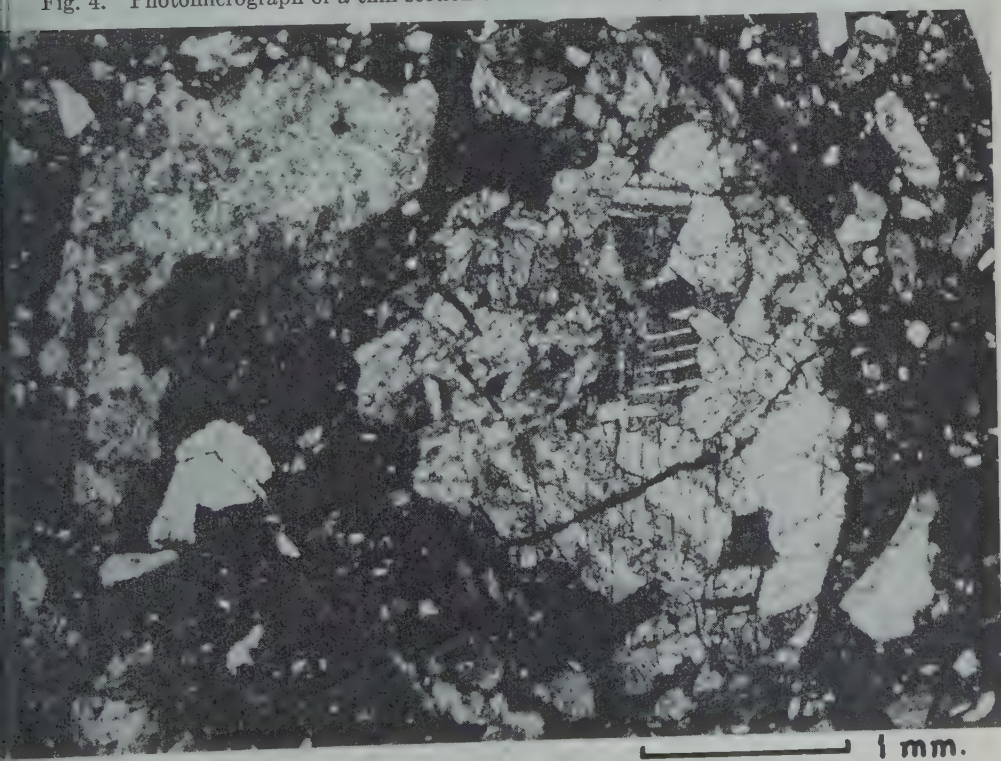


Fig. 5. Photomicrograph of a thin section of suevite from Zipplingen (crossed nicols).

brown in thin sections. It is chiefly glass with numerous trichites or hairlike crystallites. The glass has an index of refraction of 1.534, and the trichites are much higher in refringence. X-ray study shows that the trichites are crystallites of a monoclinic pyroxene probably isostructural with aegirine. The pyroxene has not yet been identified. The only X-ray pattern that is nearly identical with it is a pyroxene yet to be described from sintered siliceous dolomite from Meteor Crater, Arizona. Some low-index glass is observed as inclusions in the grayish brown glass with the trichites.

Suevite exposed in a road north of Zipplingen, on the opposite rim of the crater, contains a far larger proportion of unsintered crystalline rocks than that at Otting. Many of the larger crystalline fragments are minutely shattered, but granitic fragments tested did not contain coesite. The matrix of the breccia is characterized by a wide variety of crystalline fragments (Figs. 4 and 5). In the thin section illustrated, the larger fragment is a hornblende diorite. To the left is

a fragment with garnet and pinitic chlorite which could be the alteration product of coesite. The other poorly sorted, angular, shattered fragments are quartz, feldspar, and biotite. At least three types of glass are present: a pale brown glass with an index of 1.532; a clear glass with an index of 1.472; and a pale yellowish glass with an index of about 1.46. Some of the clear glass occurs as minute spindles some with twisted or hooked ends (Fig. 6); the brownish glass is more massive but vesicular. Quartz grains in one of the fractured granitic gneiss fragments examined for coesite are broken by closely parallel fractures perpendicular to the grain elongation. This is a common feature of quartz grains in fractured quartzose Coconino sandstone of Meteor Crater, Arizona.

Bubbles of white glass with an index of about 1.47 occur in the suevite from Zipplingen (Fig. 7). They resemble puffed rice and are exceedingly fragile.

A quarry at Aufhausen, south of the Ries, is developed in suevite very similar to that at

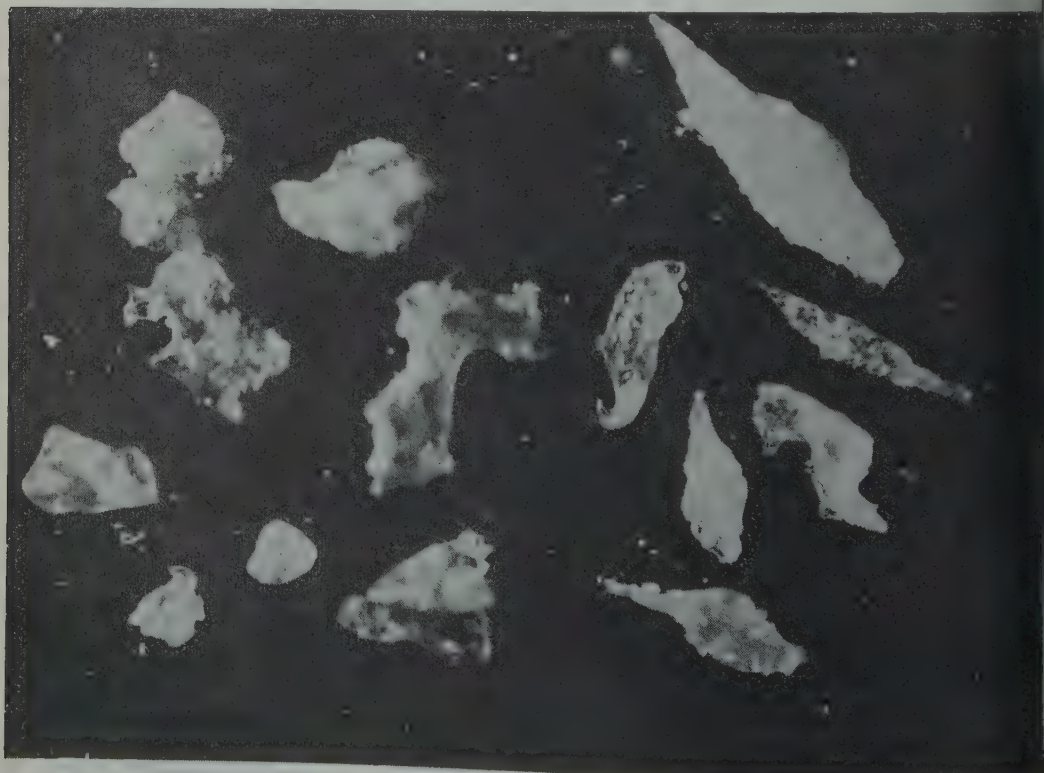


Fig. 6. Photomicrograph of spindly glass fragments from suevite.

The suevite is nearly massive and has apilli and peculiarly curved sheetlike of dark-colored glass. A fragment of white from the suevite was found to contain abundant harmotome. In a quarry ding, near by, considerable dark glass common on the floor of the quarry but much higher in the walls.

Stenburg, within the Ries crater, the exposed in a quarry is nearly massive with a vague subhorizontal jointing, as at other localities, which may be nothing more than weathering formed during weathering. Along quarry wall the suevite has a very steep to vertical contact with a giant breccia of limestone blocks.

Chemical analyses of suevite [Oberdorfer, 1960] show that the silica content ranges from approximately 56 to 68 per cent, alumina from 10 to 15 per cent, total iron from 4 to 6 per cent, and alkalis from 4 to 9 per cent. Schowalter [1903] considered the suevite to be closely related to dacite or trachyte, but the opinion merely reflects that of the various types and partly sintered crystalline rocks

of which it is composed. The minutely fractured condition of much of the fragmental material and the presence of coesite, lechatelierite, and other glasses of various composition in a single small specimen all suggest that the fusion was due to strong shock.

The suevite has been commonly assumed to have erupted from numerous widely scattered vents, and, because the suevite can locally be seen to rest on the other breccias, the eruptions have been interpreted as one of the latest events in the development of the Ries. On the basis of about a week's field study, Shoemaker believes that the suevite can be interpreted to rest everywhere on the other breccias; local steep contacts of suevite with other breccia within the crater are probably due to faulting or inward slumping of the breccias along crater walls. Rare dike-like bodies of suevite that have been observed may be crevasse fillings of some kind. The patches of suevite that have been preserved from erosion appear to be remnants of a layer that is analogous to a layer of mixed debris interpreted as fallout that is preserved in Meteor Crater, Arizona [Shoemaker, 1960]. The fallout at Meteor Crater

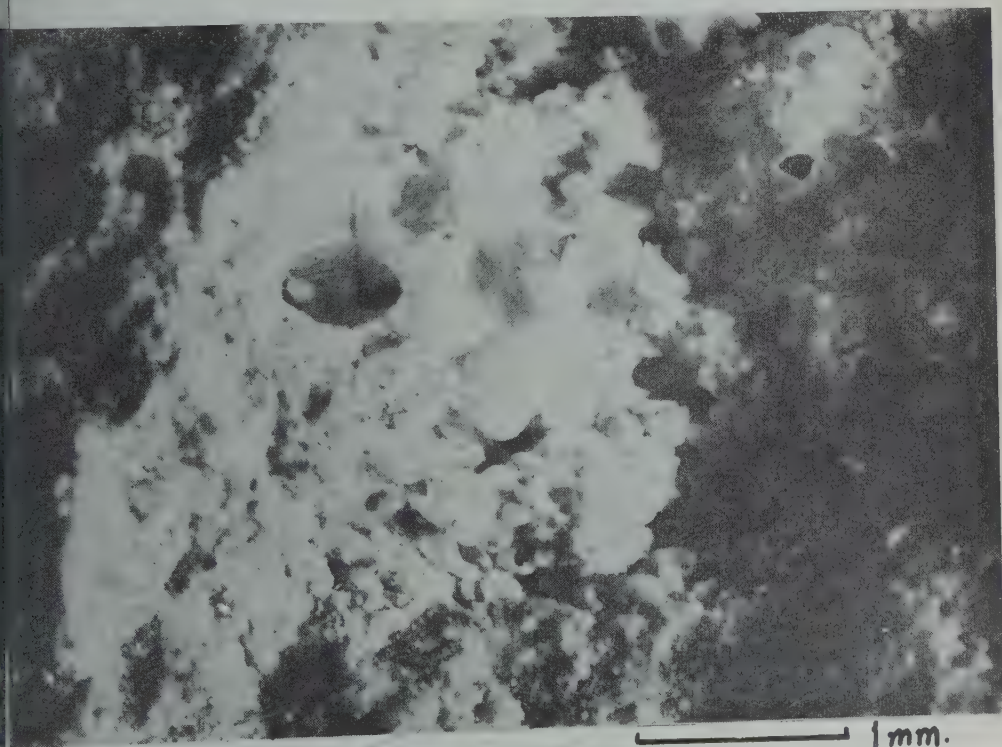


Fig. 7. Photomicrograph showing glassy bubbles in suevite.

also fills crevices in underlying breccia, where the contacts are observed in a shaft in the crater floor.

Structurally the Ries is utterly unlike any caldera or crater of demonstrable volcanic origin. Indeed, suevite is the only material at the Ries that remotely resembles volcanic rock. Kranz [1911, 1934] has attempted to explain the Ries in terms of one large volcanic explosion, but, as Reck has pointed out [Williams, 1941, p. 303], no masses of rock even approaching the size of the thrust slices of the Schollen- and Schuppen-Zone have ever been ejected in the most violent historic volcanic eruptions. To explain the thrust slices, various complicated hypotheses have been invented that involve first an uplift or doming of the central part of the Ries and later a subsidence. All the major structural features of the crater and the ejecta, on the other hand, appear to have a straightforward explanation in terms of hypervelocity impact mechanics [Shoemaker, in press].

REFERENCES

- Bentz, Alfred, Die Entstehung der 'Bunter Breccie,' das zentral Problem im Nördlinger Ries und Steinheimer Becken, *Centr. Mineral., B*, pp. 97-104, 141-145, 1925.
- Bentz, Alfred, Geologische Beobachtungen am westlichen Riesrand, *Z. deut. geol. Ges.*, 79, 405-438, Stuttgart, 1927.
- Branco, Wilhelm, Das vulkanische Vorries und seine Beziehung zum vulkanischen Ries bei Nördlingen, *Abhandl. kgl. preuss. Akad. Wiss. Berlin*, 1902.
- Chao, E. C. T., E. M. Shoemaker, and B. M. Madsen, First natural occurrence of coesite, *Science*, 132(3421), 220-222, 1960.
- Cotta, Bernhard, Geognostische Beobachtungen im Riesgau und dessen Umgebungen, *Neues Jahrb. Mineral. Geognosie, Geol. u. Petrefaktenk.*, pp. 307-318, 1834.
- Dehm, Richard, Geologische Untersuchungen im Ries—Das Gebiet des Blattes Monheim, *Neues Jahrb. Mineral., Geol. u. Paläontol., B.-Bd.*, 67 (sec. B), 139-256, 1932.
- Dorn, Paul, Ein Jahrhundert Riesgeologie, *Z. deut. geol. Ges.*, 100, 348-365, Berlin, 1948.
- Gerstlauer, K., Geologische Untersuchungen im Ries; Das Gebiet des Blattes Offingen, *Abhandl. bayer. Oberbergamt geol. Landesuntersuchung*, 1935, Munich, 1940.
- Gümbel, C. W., Über den Riesvulkan, *Sitzb. Akad. Wiss. München*, 1870.
- Knebel, Walther, Weitere geologische Beobachtungen am vulkanischen Ries bei Nördlingen, *deut. geol. Ges.*, 55, 23-44, 1903.
- Kranz, Walter, Das Nördlinger Riesproblem, *Oberrheinische geol. Ver., Stuttgart, Jahresber. Mitt.*, N. F., 1, 32-35, 1911.
- Kranz, Walter, Fünfte Fortsetzung der Beiträge zum Nördlinger Ries-Problem, *Centr. Mineral.*, no. 6, pp. 262-271, 1934.
- Nathan, Hans, Geologische Untersuchungen im Ries—Das Gebiet des Blattes Möttingen, *Neues Jahrb. Mineral., Geol. u. Paläontol., B.-Bd.*, (sec. B), 31-97, 1925.
- Nathan, Hans, Geologische Untersuchungen im Ries—Das Gebiet des Blattes Ederheim, *Abhandl. bayer. Oberbergamt geol. Landesuntersuchung*, no. 19, 42 pp., Munich, 1935.
- Oberdorfer, R., Die vulkanischen Tuffe des Ries bei Nördlingen, *Jahrb. Ver. Naturk. Würtb.*, pp. 1-40, 1905.
- Reich, Hermann, and Wilhelm Horrix, Geophysikalische Untersuchungen im Ries und Vorries und deren geologische Deutung, *Geol. Jahrb., Beih.*, no. 19, 119 pp., Hanover, 1955.
- Reuter, Lothar, Die Verbreitung jurassischer Kalkblöcke aus dem Ries im sudbayr. Diluvialgebiet, *Oberrhein. geol. Ver., Stuttgart, Jahresber. u. Mitt.*, N. F., 14, 191-218, 1925.
- Sauer, Adolf, Petrographische Studien an den Lavabrocken aus dem Ries, *Jahrb. Ver. Naturk. Würtb.*, 57, 1901.
- Schröder, Joachim, and Richard Dehm, Geologische Untersuchungen im Ries, *Abhandl. Naturw. Ver. f. Schwaben e. V. Augsburg*, 5, 147 pp., 1950.
- Shoemaker, E. M., 1960, Penetration mechanics of high velocity meteorites, illustrated by Meteor Crater, Arizona, *Rept. 21st Intern. Geol. Congress Copenhagen*, pt. 18, pp. 418-434, 1960.
- Shoemaker, E. M., Geologic interpretation of lunar craters, in *The Moon, Its Astronomy and Physics*, edited by Zdenek Kopal, Academic Press, New York, in press.
- Trieb, Walter, Geologische Untersuchungen im Ries; das Gebiet des Blattes Otting, *Geol. Bavarica*, no. 3, 52 pp., Munich, 1950.
- Werner, E., Das Ries in der Schwäb.-fränk. Alb., *Blätter Schwäb. Albvereins*, 1904.
- Williams, Howel, Calderas and their origin, *Univ. Calif. Dept. Geol. Sci. Bull.*, 25(6), 239-346, 1941.

(Manuscript received July 10, 1961.)

Analysis of the Formation of Meteor Crater, Arizona: A Preliminary Report¹

R. L. BJORK

*Aero-Astronautics Department, The Rand Corporation
Santa Monica, California*

Abstract. A theoretical study is made of the cratering process accompanying the impact of a 100-ton iron projectile on a semi-infinite half-space of soft rock at a velocity of 30 km/sec. The constituents and velocity approximate those involved in the formation of Meteor Crater, Arizona. The assumption is made that the process is hydrodynamic in nature, since the pressures generated so greatly exceed the strengths of the materials. At these high pressures, the compressibilities of the materials must be taken into account, with the result that shocks are generated. The motion is solved by numerical means, and graphs showing details of the motion are presented. The conclusion in this preliminary report is that the meteorite had a mass between 30,000 and 194,000 tons, the range being due to the uncertainty in the impact velocity.

Introduction. The Arizona Meteor Crater, in the Canyon Diablo region of north-central Arizona, is a well-known geological feature. Its very name calls up visions of the fantastic violence that must have accompanied its creation. Its depth, about 570 feet from rim to bottom, and its diameter of approximately 4000 feet, are impressive. Details of the geometry and geology of the crater and its environs are available in a report by Shoemaker [1960], who also gives an extensive bibliography on the subject. Remnants of the meteorite itself are scattered throughout the crater. By far the greatest mass of meteoritic material is in the form of small iron particles distributed throughout the soil within a few miles of the crater. In a 1956 expedition sponsored by the Smithsonian Institution, Rinehart [1958] made a careful sampling of these, concluding that about 12,000 tons of meteoritic iron is present in this form, the majority of particles being between 0.5 and 2 mm in dimension. In a later report, Rinehart cites that between 20 and 100 tons of larger meteorites had previously been recovered from the immediate vicinity, these ranging in size from a few ounces to about 2000

tons. In attempting to analyze the creation of the crater, one is immediately confronted with two major uncertainties: the mass and the impact

velocity of the meteorite. The evidence is fairly conclusive that the meteorite was composed mainly of iron and had a mass of at least 12,000 tons. It is also extremely likely that its impact velocity was between 11 and 72 km/sec. Of the thousands of velocity measurements made on meteorites, all lie in this range, which corresponds to earth-escape velocity on the one hand and the maximum velocity any member of our solar system could have with respect to the earth on the other. Assuming that it had the average meteor velocity of 30 km/sec could thus lead to an error of at most a factor of 2.7. Previous studies on hypervelocity impact lead one to suppose that the mass deduced from such an assumption is in error by the same factor as the velocity [Bjork, 1958].

The possible velocity error is small compared with the range of previous estimates of the meteorite mass, which extends from the 10,000–15,000 tons estimated by Rinehart [1950] to the 5,000,000 tons estimated by Öpik [1936] and later by Rostoker [1953].

It is unfortunate that impact velocities of meteoric magnitude have not so far been achieved in the laboratory under circumstances that would permit the making of quantitative measurements on the crater produced. However, some fairly recent results have been obtained that make it possible to deduce that the mass estimate of 5,000,000 tons is substantially too high. An iron sphere of this mass would have a diameter of about 100 meters, and the implication would be

¹Presented at the Geophysical Laboratory–Space Radiation Laboratory Cratering Symposium, held at the Geophysical Laboratory in Washington, D. C., on March 28 and 29, 1961.

that in striking the earth at at least 11 km/sec it produced a crater only 2 sphere diameters in depth. Experiments have shown that iron spheres at the much lower velocity of 6 km/sec produce craters of about 2 sphere diameters' depth even in *steel* targets (W. S. Partridge, Utah Research and Development Corp., Salt Lake City, personal communication).

Model. To simulate the creation of the Arizona Meteor Crater, we consider a 12,000-ton iron projectile striking a semi-infinite target of tuff, a soft rock. The impact velocity is taken to be 30 km/sec directed at right angles to the ground's surface. The whole problem has axial symmetry if the projectile does. Accordingly, the geometry chosen for the projectile is a right circular cylinder having the same length as its diameter. For the mass chosen, both dimensions turn out to be 12 meters. The cylinder axis is in the direction of the initial motion.

Hypervelocity impact experiments have shown that the size and shape of the crater are not strongly dependent on the geometry of the projectile as long as it is not extreme, such as a long, pencil-shaped jet traveling along its axis. If all dimensions of the projectile are of the same order, it does not matter greatly whether it is spherical, cubical, cylindrical, or irregular.

Therefore, the material and geometry of the projectile are thought to be well chosen.

The target is not so well approximated. The actual target structure consists of strata, roughly parallel to the ground surface, whose members are a 270-foot layer of Kaibab limestone underlain by a 700- to 800-foot slab of Coconino sandstone [Shoemaker, 1960]. Tuffaceous porous volcanic rock, is less dense and somewhat softer than these materials, and is used only because its equation of state was available to the author, having been worked out by Brode in connection with another problem [Brode and Bjork, 1960].

The velocity provides another area of uncertainty, but fortunately fairly concise limits may be placed upon it. Of the thousands of meteor velocities measured, all lie between 11 and 72 km/sec. The lower limit corresponds to the earth-escape velocity; the upper limit is the greatest velocity any member of our solar system could have with respect to the earth. The average of measured meteor velocities is about 30 km/sec. Choosing it means that the error is at most a factor of 2.7 and probably is much less.

The angle of incidence is also unknown. However, experiments have shown that the crater dimensions depend only on the normal

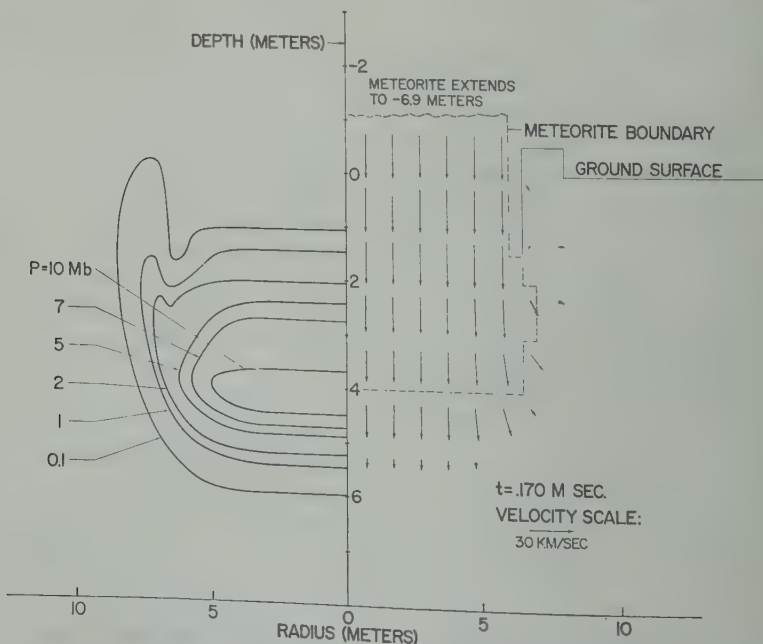


Fig. 1. Pressure and velocity field at 0.17 msec.

nt of velocity up to angles of incidence
55° [Summers and Charters, 1959], so
case we are treating is that of a meteorite
a normal velocity component of 30

regime of pressures generated by impact
tric velocities, materials behave in ways
to most of our intuitions. We are used
ing of massive iron as a very strong
e, which under the most extreme forces
slightly distorted. But under the condi-
e are considering, the massive iron
is immediately squeezed into only
original volume by the pressure generated
ct. The softer tuff is compressed to less
urth of its normal volume. This feat is
shed by a pressure of 10.5 megabars
(about a million atmospheres).

light of this information it is under-
e why the strength of both the iron and
y be neglected in calculating the motion
by the impact. The pressure exceeds the
y by such a vast amount that any pressure
s will cause the material to literally
a fluid. Therefore the approximation is
that the compressible hydrodynamic
s govern the motion.

is framework, the material properties
cribed by an equation of state which
the pressure as a function of density
ernal energy. Only pressure-type forces
sidered, since the material has insufficient
h to support any shear comparable to
sure. In arriving at the equation of state
ver the requisite pressure range, two
of information are used. At pressures
about 10 Mb, the quantum-mechanical
Thomas-Dirac theory is thought to be
Experimental data are available between
0 and 300 kb for tuff, and between about
5000 kb for iron. These data are obtained
g high explosives to generate shocks in
materials under consideration. The equations
these calculations are: for tuff

$$.425\eta e + 0.113\eta^{3/2}e + 5.30\eta e^{1/2} + \frac{0.707\eta e^2}{(10^5 + e)} \quad (1)$$

iron

$$P = \frac{a_1\mu + a_2\mu|\mu| + \epsilon(b_0 + b_1\mu + b_2\mu^2) + \epsilon^2(c_0 + c_1\mu)}{\epsilon + \epsilon_0} \quad (2)$$

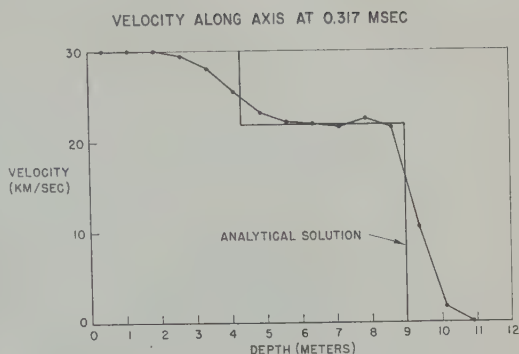


Fig. 2. Velocity along axis at 0.317 msec.

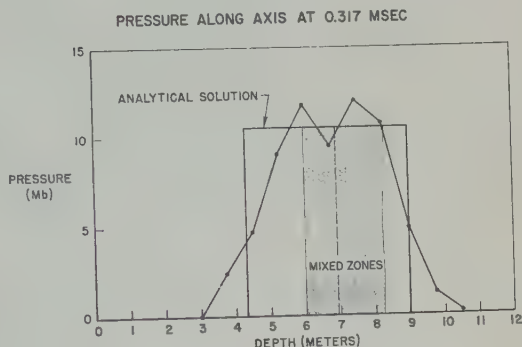


Fig. 3. Pressure along axis at 0.317 msec.

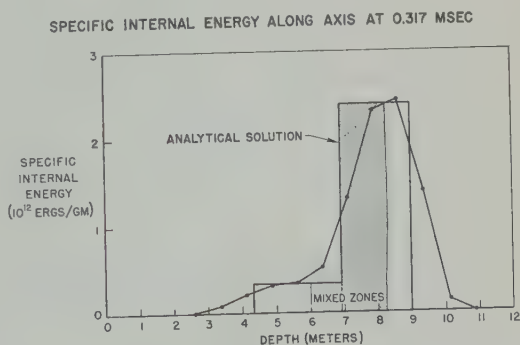


Fig. 4. Specific internal energy along axis at 0.317 msec.

(This equation of state was furnished by F. Harlow of the Los Alamos Scientific Laboratories. The constants were determined by Osborne and associates in group W-4.)

The units for equations 1 and 2 are the

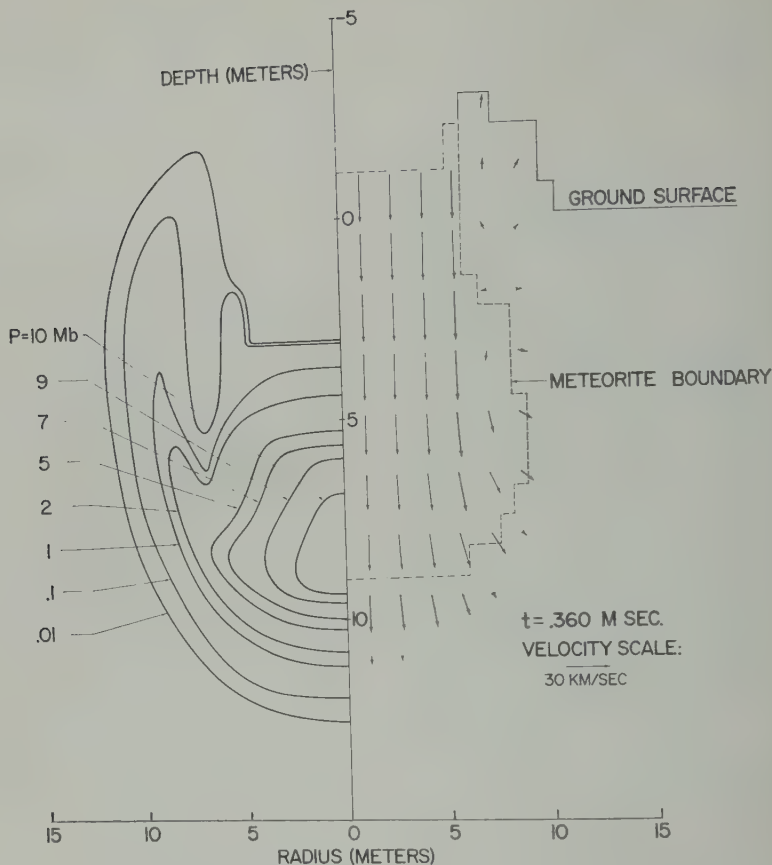


Fig. 5. Pressure and velocity field at 0.360 msec.

megagram, millisecond, meter system, and the symbols are

ρ = density.

P = pressure.

ρ_0 = normal density (1.7 for tuff, 7.86 for iron).

$\eta = \rho/\rho_0$.

$\mu = \eta - 1$.

e = specific internal energy.

$\epsilon = \rho_0 e$.

The values of the constants are

$$a_1 = 7.780 \times 10^4. \quad b_2 = 463.4.$$

$$a_2 = 31.18 \times 10^4. \quad c_0 = 0.3984.$$

$$b_0 = 959.1. \quad c_1 = 0.5306.$$

$$b_1 = 1568. \quad \epsilon_0 = 900.$$

The initial conditions for the problem are set at the instant the cylindrical projectile first contacts the ground. At this time, the pressure and internal energies are everywhere zero, and all materials have their normal density. All the meteorite material is given a velocity of 3 km/sec. These are the only inputs to the problem besides the constants specifying the equation of state.

Subject to these initial conditions a numerical solution of the compressible hydrodynamic equations is generated on an IBM 7090 computer [Bjork and Brooks, 1960]. During the calculation the boundary condition that the pressure at solid-vacuum interface is zero is imposed. The results are discussed in the following section.

Results. Figure 1 illustrates conditions 0.1 millisecond after the initial contact between the meteorite and the ground. The original cylindrical

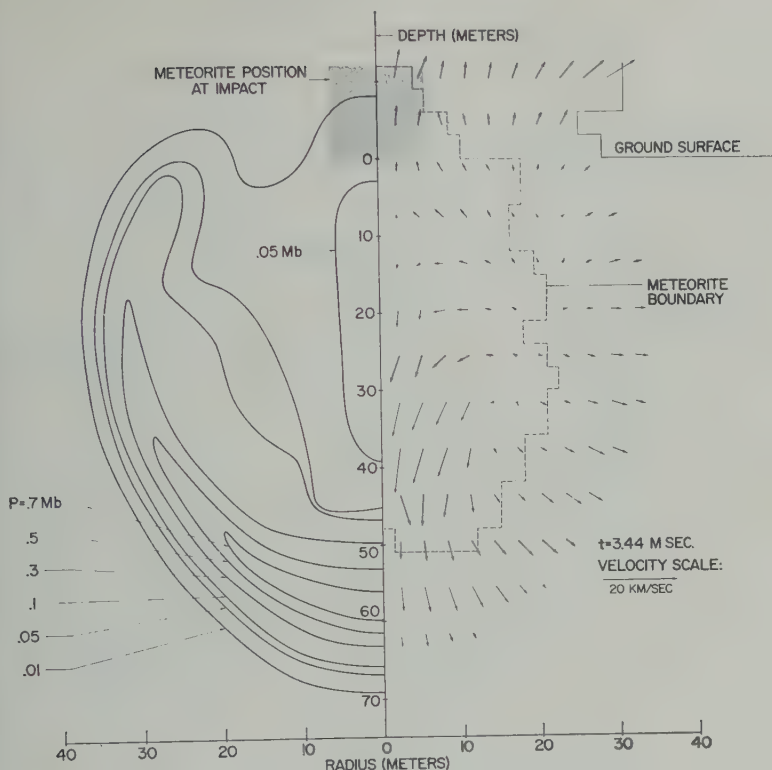


Fig. 6. Pressure and velocity field at 3.44 msec.

had a radius of 6 meters and a length of 10 meters. The y axis corresponds with the axis of symmetry of the problem, the velocity field is presented on the right and the pressure field is shown by the contours on the left. Each vector represents the velocity of the particle situated at the tip of the vector. For clarity, only one-half of the available velocity data are presented in every other row and column having been included. It is clear that, near the axis of symmetry, the velocity vectors remain parallel to the y axis, so that the flow is still one-dimensional. For such a flow analytical solutions are available, and it is possible to compare the numerical solution with them to see how well the numerical solution is doing.

The numerical velocities are compared in Figure 2. The numerical solution predicts the presence of two shock surfaces across which discontinuities in pressure, density, internal energy, and velocity occur. The projectile material moves unimpeded at its original velocity of 30 km/sec until it reaches the upper shock, shown as the left one

in the figure. At the time of 0.317 msec the shock is at a depth of slightly more than 4 meters. The projectile material is slowed discontinuously to a velocity of 21.84 km/sec as it crosses the shock, being simultaneously compressed and brought to a high pressure and internal energy. Similarly, the target material remains at rest until the leading shock reaches it. At 0.317 msec, the leading shock is at a depth of almost exactly 9 meters. As the shock engulfs it, the target material is suddenly given a velocity of 21.84 km/sec, the same as the shocked projectile material. The continuity of material velocity across the interface between the two materials (which at this time is at about 7 meters) is always preserved in the one-dimensional case.

The numerical solution, shown by the broken line in Figure 2, does not match the analytical solution exactly. The most salient feature is that the velocity increase does not take place discontinuously, but requires on the order of three grid spaces to occur. This feature is required in

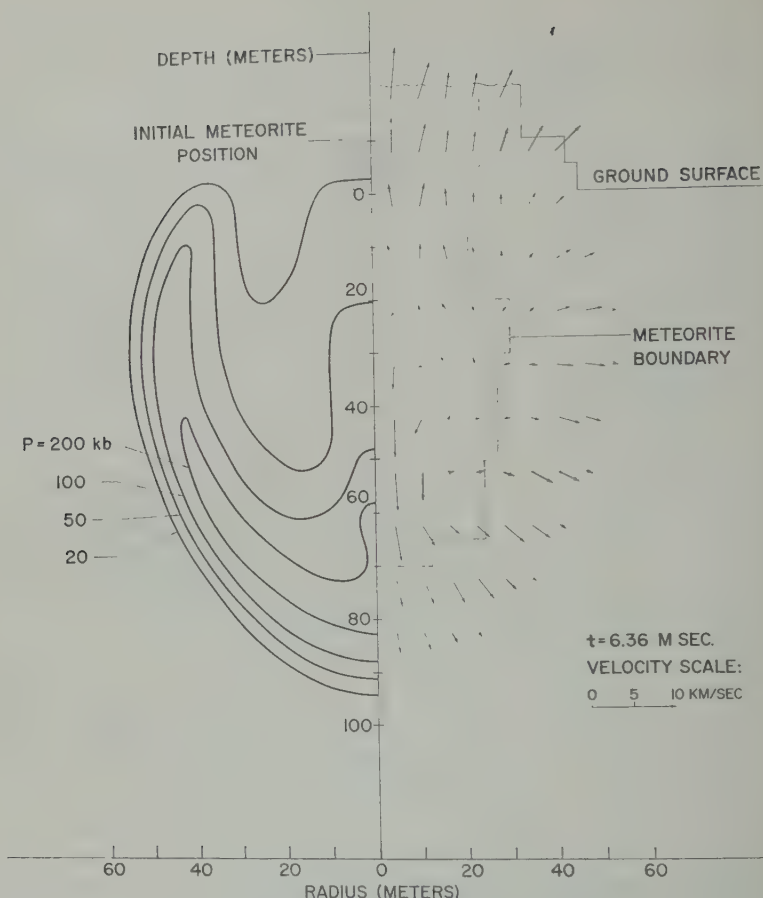


Fig. 7. Pressure and velocity field at 6.36 msec.

order for the numerical solution to be stable. However, it is seen that the velocity increments are nearly correct, and that the transition corresponds well with the position predicted for the shocks.

Figure 3 compares the pressures. In the analytical solution there is no pressure outside the shocked region, emphasizing that the projectile and target material remain in their initial velocity state until a shock is encountered. A uniform pressure of 10.5 Mb is predicted for the shocked material. Again, the numerical solution rises rapidly, but continuously to simulate the shock, the transition occurring about where the shocks should be, and approximately the correct pressures being obtained behind the shocks. Oscillations are seen to occur in the numerical solution. Following the pressure

history of an individual particle, we see that the pressure oscillates in time about the correct value, the average pressure being such as to maintain the correct shock and particle velocity.

Specific internal energies are compared in Figure 4. Both the iron and the tuff are brought to the same pressure behind the shock, but the tuff undergoes a far greater volume change. We can see intuitively that the $P dV$ work done on it must therefore be far greater than in the case of iron. In addition to this effect, there is the smaller mass of tuff per unit volume. Both factors contribute to tuff's acquiring a greater specific internal energy than iron in the shock compression.

Behind the shock the density of the iron is 1.99 normal and that of the tuff is 4.38 normal. Referring back to Figure 1, we see that

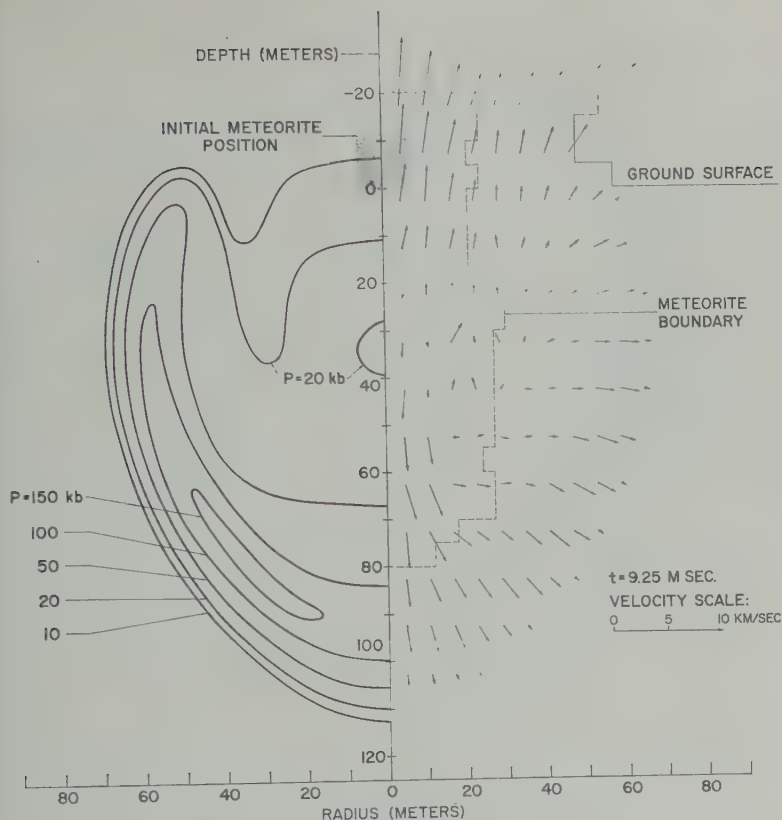


Fig. 8. Pressure and velocity field at 9.25 msec.

interface is at about 4 meters' depth. The dashed line is the boundary between iron and tuff. Since it is clear from the velocity field that there is a slight amount of lateral spreading of the meteorite at this time, a 4-meter cylinder of iron has been wadded into a 1-meter cylinder, between about 4 and 5 meters' depth at this time.

At this cylindrical hole which has been formed in the tuff, the ground has recoiled and is generating a pressure field described by the wings in the 0.1- and 1-Mb contours. Since the acceleration is proportional to the pressure gradient, we can see the reason why the tuff starts to 'rebound' into the zero pressure region corresponding to the space between the side of the meteorite and the tuff, and the upward acceleration of the tuff into the region above the ground surface.

The nature of the motion continues, as shown in Figure 5. Here the ground has rebounded into the region of the meteorite and more of it is being

hurled into the air. More lateral spreading of the meteorite has occurred, with the result that the high-pressure contours are shrinking as rarefaction waves proceed into them. The meteorite material is proceeding unimpeded below the ground surface level until it encounters the upper shock, which at this time of 0.36 msec is at a depth of about 5 meters.

By 3.44 msec, as shown in Figure 6, extensive mixing of the iron and tuff has occurred. The front portions of the meteorite are continually peeling off and being driven laterally into the ground. Then, as the ground rebounds toward the axis, they are carried back. As the rebounding material reaches the axis, it builds up a high pressure and is deflected upward or downward. What might be termed a stagnation point exists at a depth of about 12 meters, material above this point being deflected upward, and below it being deflected downward to follow the meteorite. In the machine program, three types of regions are recognized: pure iron, pure tuff, and a mixture.

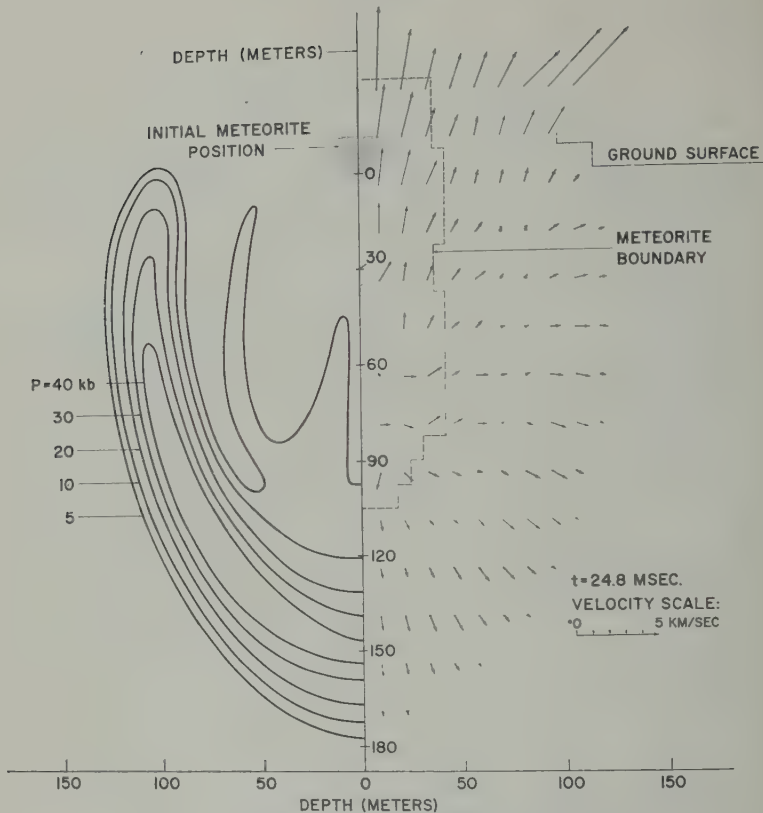


Fig. 9. Pressure and velocity field at 24.8 msec.

At this time only two types exist, pure tuff and mixture, the dashed line denoting the interface of the two. The meteorite material is spread throughout many times its original volume.

In Figures 7 and 8 the mixing motion is seen to continue, the stagnation point moving lower and lower, so that an increasing amount of material is hurled upward. Another important feature of the motion during this time period is the shock which is being driven into the tuff and the flow induced behind it which occupies an ever-increasing volume.

In Figures 9 and 10 the mixing has essentially stopped, and the flow in the tuff has been generated which will lead to the crater. It is estimated in this preliminary report that the final crater will be about 150 meters deep and have a radius of about 500 meters. In Figure 10 it is seen that intense shocks penetrate substantially deeper than the crater bottom, and these are believed to account for the brecciated region beneath the crater cited by Shoemaker.

Conclusions. The solution indicates that the

meteorite will be thoroughly pulverized and mixed with the ground material, so that no portions of it are expected to be found intact.

The meteorite's energy is delivered into the ground across the surface of a deep, relatively narrow hole, acting more like a line source than a point source. The top portions of the line are loaded first, and then the force is released. The ground rebounds to the axis, and the convergent flow generates additional shocks which then proceed radially outward. This complicated flow is very different from that generated by an explosive detonated at a point. Consequently, descriptions of the crater formation modeled by explosive experiments are likely to be in considerable error.

A preliminary estimate of the theoretical crater's size as given by these calculations is a depth of 150 meters and a radius of 500 meters. These dimensions are about 20 per cent smaller than those of the actual crater. Since the theoretical results scale with the meteorite dimensions, the two may be brought into correspondence

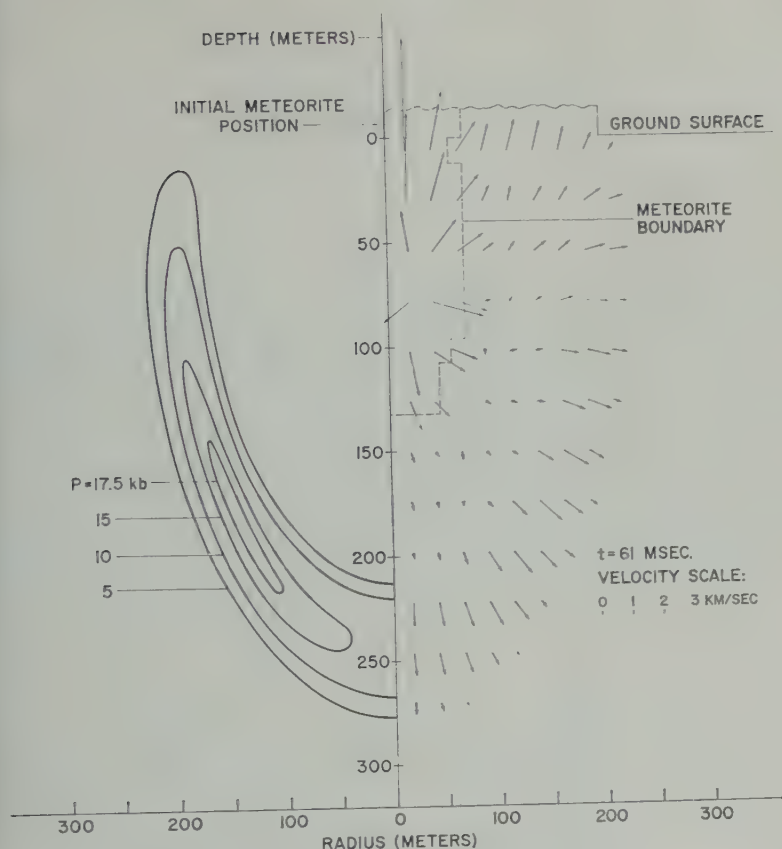


Fig. 10. Pressure and velocity field at 61 msec.

that the meteorite had a length and mass of about 21,000 tons.

It is difficult to estimate the crater size that would be calculated if limestone and sandstone were used instead of tuff, but a rough estimate is that the crater calculated here is about 50 per cent larger than would be obtained with these materials. This would indicate that a body of diameter 21.6 meters would be required to produce the crater in limestone and sandstone, and such a meteorite would have a mass of about 71,000 tons.

If the impact velocity were 11 km/sec, the mass would be 194,000 tons; if it were 15 km/sec, about 30,000 tons.

REFERENCES

L., Effects of a meteoroid impact on aluminum in space, *Rand Corp. Paper* December 16, 1958.

- Bjork, R. L., and N. B. Brooks, A numerical technique for the solution of multidimensional hydrodynamic problems, *Rand Corp. Research Mem. RM-2628*, August 10, 1960.
- Brode, H. L., and R. L. Bjork, Cratering from a megaton surface burst, *Rand Corp. Research Mem. RM-2600*, June 30, 1960.
- Öpik, E., *Publ. de l'Observatoire astron. Univ. Tartu*, 28, 27 pp., 1936.
- Rinehart, J. S., *Popular Astron.*, 58, 458, 1950.
- Rinehart, J. S., *Smithsonian Contrib. to Astrophys.*, 2, 145-160, 1958.
- Rostoker, N., *Meteoritics*, 1, 11, 1953.
- Shoemaker, E. M., Penetration mechanics of high velocity meteorites, illustrated by Meteor Crater, Arizona, *Rept. Intern. Geol. Congr., XXI Session*, Norden, 1960, Copenhagen, 1960.
- Summers, J. L., and A. C. Charters, *Proceedings of the Third Symposium on Hypervelocity Impact*, edited by F. Genevieve, Armour Research Foundation of Illinois Institute of Technology, Chicago, Ill., p. 101, 1959.

(Manuscript received July 10, 1961)

High-Explosive Craters in Desert Alluvium, Tuff, and Basalt¹

B. F. MURPHEY AND L. J. VORTMAN

Sandia Laboratory, Albuquerque, New Mexico

Abstract. Explosion craters in desert alluvium have been formed over a range of energy releases from 256 to 1,000,000 pounds of TNT. An empirical scaling law in which crater dimensions vary with energy release to the 0.3 power best relates dimensions from small to large chemical explosions. Consideration of overburden leads to a partial explanation of the failure of cube-root scaling. Optimum depth of burst for 256-pound chemical explosions in desert alluvium is near 1.75 times maximum depth. Thirteen 256-pound charges of spherically cast TNT were detonated in desert alluvium to determine apparent crater dimensions. Charges were placed at six different burst depths in the region approaching containment of the explosion. Variation of crater dimensions with burst depth was determined. No crater resulted where burst depth was greater than 1.75 times maximum depth. Ten 1000-pound charges (two at each of five burst depths) and three 40,000-pound charges (one at different burst depths) describe variation of crater dimensions with burst depth in basalt. At constant charge size, maximum and average rock size increases with increased burst depth. At constant scaled burst depth, maximum rock size increases as charge weight is increased.

Introduction. Craters produced by underground explosions have dimensions which depend on several parameters, the most obvious of which are (1) explosive energy release, (2) depth of burst, and (3) type of earth in which the explosion takes place. With the advent of the Plowshare program, a program to study industrial applications of nuclear explosives, it became necessary for engineering purposes more accurate data were required than could be found in the literature [Waterways Experiment Station, Vicksburg, Miss.]. Existing craters produced by nuclear explosives occur in desert alluvium and tuff at the Nevada Test Site. A few water-washed craters produced by approximately surface bursts from nuclear explosions exist at the AEC Pacific Proving Grounds.

Crater studies reported here represent an attempt to improve our knowledge of craters produced by chemical explosives. Another objective was to relate craters from chemical explosives to nuclear explosives; hence, much of the work was done in desert alluvium. The range of energy releases employed was from 256 to 1,000,000

pounds of TNT, the latter of which was the Scooter shot. Spherically shaped charges were used, either actual cast spheres for 1000 pounds and less or blocks stacked in a spherical shape for larger charges. All charges were centrally detonated.

Yield and depth-of-burst dependence in desert alluvium. Desert alluvium is the material that fills the valley floor of Yucca Flat (and similar flats in the West) exclusive of the dry lake bed (playa). Texture of the alluvium is variable, including the whole range of materials from clay and silt-sized particles to cobbles and boulders, lightly cemented. Density is near 1.5 g/cc. Strength of the material depends on the amount of overburden.

The half-kiloton (i.e., the 1,000,000-pound Scooter shot) charge was buried at a depth of 125 feet. The depth of the crater shown in Figure 1 is 75 feet; the diameter at original ground level is 307 feet. The flat bottom is very soft, powdery material and lies 50 feet above the original position of the charge center. Another view of the crater, Figure 2, gives some impression of the lip size, which was only 8 or 9 feet. This crater is not quite as large as would be predicted by simple cube-root scaling from crater dimensions obtained from much smaller charges.

Presented at the Geophysical Laboratory-National Radiation Laboratory Cratering Symposium, held at the Geophysical Laboratory in Washington, D. C., on March 28 and 29, 1961.

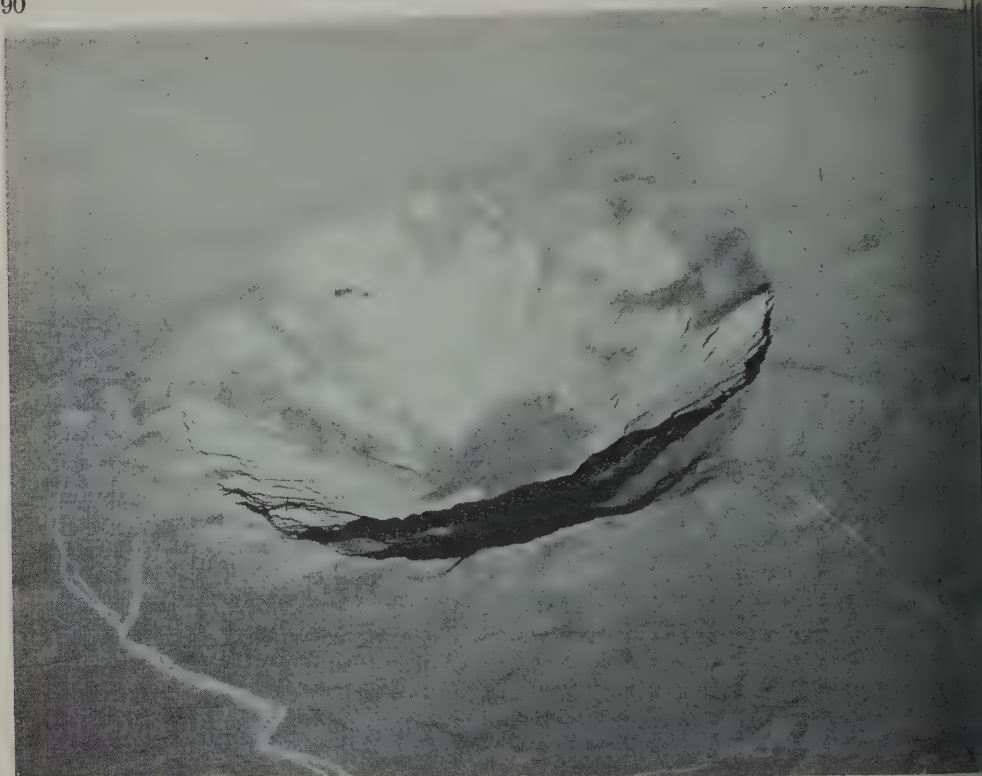


Fig. 1. Aerial view of Scooter crater.

Cross sections of craters produced by 256-pound TNT charges [Sachs and Swift, 1955; Murphey and MacDougall, 1959] burst at several depths are shown in Figure 3. The very great importance of deep, but not too deep, burial is apparent. Incidentally, the energy available in the explosive far exceeds the work done in forming any of these craters, even when required lifting of the material is taken into account. Note that no appreciable crater is formed for depths of burst beyond $DOB = 3W^{1/3}$ (W in pounds TNT). We have scaled by cube-root scaling the crater which would be at a depth of burst of 8 feet to a crater that might be obtained from 1,000,000 pounds burst at the corresponding depth (125 feet) (Fig. 4).

Quite obviously, extrapolation of cube-root scaling to large yields is not warranted; dimensions that would be predicted are in this case too large by more than 50 per cent. This result had been forecast empirically from experiments with 256, 2560, and 40,000 pounds of TNT and from observations of craters from two nuclear

explosions. The next three illustrations set forth the data with different types of scaling.

Failure of cube-root scaling appears again in Figure 5, in which all crater data obtained in desert alluvium have been reduced in accordance with the cube root of the energy release of each explosion. Note the scatter of points for small charges. Scatter is avoided for large explosions by the simple expedient of not repeating shots. Effects of nonuniformity of medium are presumably much reduced for larger explosions because of the averaging of medium properties over large volumes. These data have been fitted by trial and error to obtain the empirical scaling law illustrated in Figure 6. The rule that $W^{0.33}$ should be used has been obtained independently by at least four different groups [Chabai, 1959; Johnson, 1960]. To what extent this rule may be extrapolated to yields larger than a kiloton is not known. One possible reason for inadequacy of cube-root scaling is, of course, failure to consider gravitation forces.

If effect of gravity [Chabai and Hankins, 19



Fig. 2. View of Scooter crater lip.

ed in scaling (Fig. 7), nearly as good a obtained as was shown in Figure 6. It therefore, that the most likely cor- is the right one.

of ground motion above buried explo- provides some information about the sm of cratering. Surface motion was aphed above each of the 40,000-pound charges in the Stagecoach series placed at of burst of 17.1, 34.1, and 80 feet [*Feigen- and Wegkamp*, 1961]. Displacement, and acceleration of surface motion en obtained from motion pictures.

that time scales of the deeper explosions ended compared with those of the shal- nes. This is also true of large vs. small ns. Corresponding times are longer in ion to cube roots of the charge weights. surface velocity vs. time for the three ound shots is illustrated in Figure 8. hat velocity decreases roughly as the

depth of burial to the 2.2 power. Corresponding accelerations are shown in Figure 9.

Acceleration of the surface decreases roughly as the fourth power of the depth of burial. Thus the 20g acceleration in evidence for the 80-foot-deep shot drops rapidly toward 1g as the observa- tion point is shifted away from surface zero toward the edge of the expected crater. Where accelerations are geometrically scaled to larger charges, accelerations at corresponding distances will be smaller as the cube root of the ratio of the charge weights. Of course, times are longer in the same proportion. Nonetheless, if scaled accelerations become less than 1g, no cratering can occur from throwout of material.

Let us, therefore, inquire in more detail about the processes that occur. Scooter can be taken as an example, since some data on underground particle motion [*Perret, Vortman, Chabai, and Reed*, to be published] are available. An accel- erometer and a velocity gage were placed at

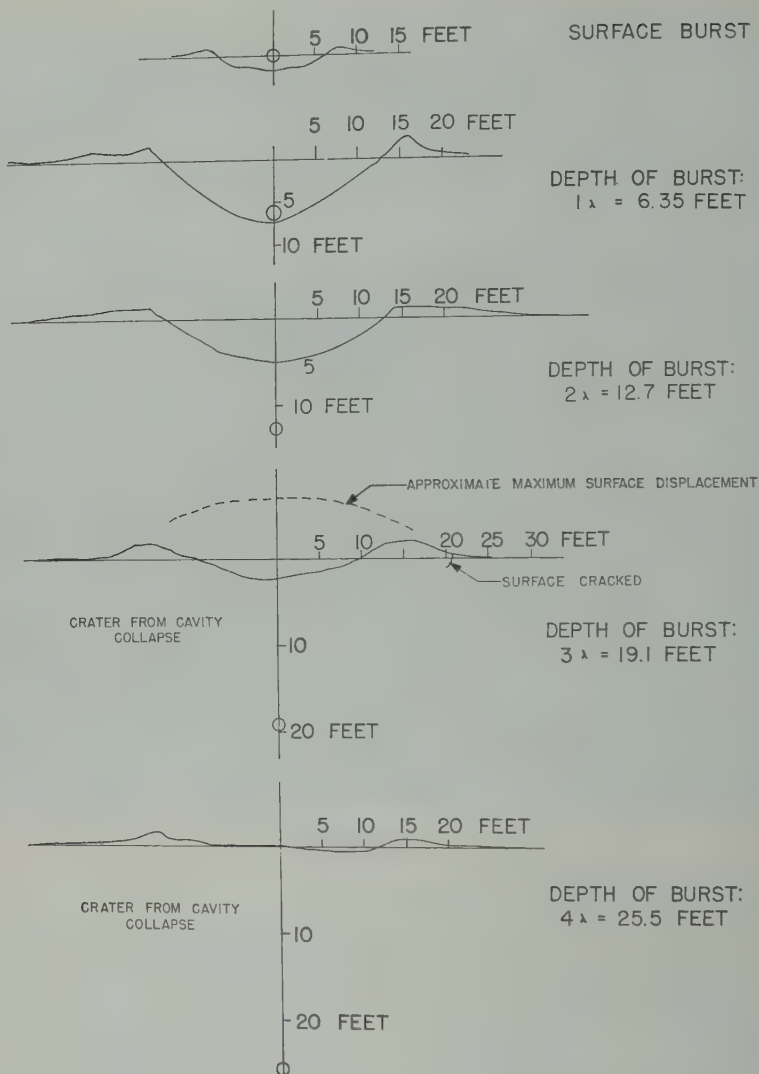
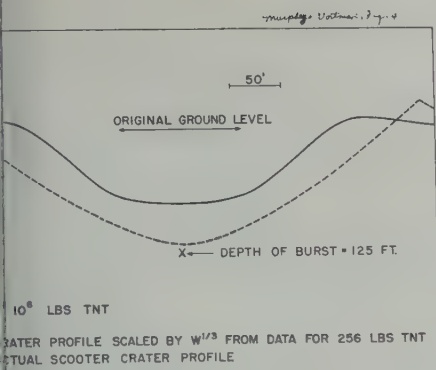


Fig. 3. Crater profiles vs. depth of burst.

shot depth on a horizontal radius 200 feet from the zero point. The distance 200 feet turns out to be equal to the slant distance at which the apparent crater radius eventually existed. The velocity pulse arrived at ± 50 m/sec and lasted for nearly 250 m/sec before changing direction. The peak value was about 8.5 ft/sec. Accelerations involved are from $2g$ to $5g$. The vertical components of peak velocity and acceleration turn out to be 5.3 ft/sec and $1.3g$ to $3.1g$, which numbers are to be doubled because of reflection from a free surface. Under the influence of gravity, free-surface upward velocity will drop

to zero in $\frac{1}{3}$ second, so that the surface has $\frac{2}{3}$ second for horizontal motion. It can go about 5 or 10 feet in this time.

Now consider an interior point where we were able to trace the motion of a target for 2.3 seconds. The target was on the surface at a horizontal distance of 120 feet from surface point above zero. Slant range was 173 feet. Interpolating from subsurface motion measurements at 150 and 200 feet, peak particle velocity in the medium was 13.1 ft/sec. The doubled vertical component is about 14.5 ft/sec. We find from analysis of motion-picture photograph



Scooter crater profile compared with scaled small charge crater profile.

early vertical velocity is 17.5 ft/sec. We find that this velocity is sustained for a second, after which it drops to zero by two seconds. Obviously something more than acquired particle motion is responsible for eventual movement. This is, of course, gas pressure and trapped momentum enlarges the original cavity in the direction of free surface. Closer to the center of the crater, surface velocity continued to decrease to the time of venting.

Displacements of four of the Scooter targets are shown in Figure 10. Venting of the black explosive explosion products started at 0.6 seconds. Thus we have some quantitative feeling for surface motions that actually took place on the crater. We have found that just beyond 150 milliseconds the motions are small. Note that on the order of 256 pounds they would still be of importance. Particle motion in the medium can be scaled from 256 to 1,000,000 pounds without departure from cube-root scaling. Failure of scaling for crater dimensions lies in the marginal region near the edge of the crater. Material starts into large motion but falls within the crater wall slumps to the bottom to fill it and decreases the apparent depth. Air drag becomes increasingly effective in reducing the distance particles travel as the time during which they act becomes longer.

Although departures from cube-root scaling are observed in desert alluvium, they have not been observed in other media. More precise data over a larger range of charge weights in other media may show that such departures do in fact

exist. Pending conclusive evidence of such departures, we have chosen to scale dimensions as the cube root of charge weight in all materials other than desert alluvium.

Dependence on tuff. During April 1959 a series of cratering shots was fired in volcanic tuff at the Nevada test site [Murphey, 1961]. This material and its properties have been described in detail [Warner and Violet, 1959]. At the site of the experiment described here density of the tuff ranges from 1.5 to 1.6 g/cm³. Its compressive strength averages about 3900 psi, with three-fourths of the measurements falling between 2000 and 6000 psi.

Charges weighing 256 pounds were placed in drilled holes in the tuff and covered with about 1 foot of sand. The remaining hole space was stemmed with concrete, the density and compressive strength of which approximated those of the tuff. The number of shots at each scaled depth of burst are tabulated below.

No.	Scaled Burst Depth, ft/ $W^{1/3}$	No.	Scaled Burst Depth, ft/ $W^{1/3}$
2	1	3	2.5
3	1.5	2	3
2	2	1	3.5

Figures 11 and 12 show, respectively, scaled apparent crater radius and depth as a function of scaled burst depth. No crater was obtained at scaled burst depths greater than 1.75 ft/ $W^{1/3}$, where W is the charge weight in pounds. At greater depths the result was a mound of rock which has been shown in Figure 12 as a negative crater depth. With further increase in depth, the height of the mound gradually decreased until, somewhere beyond 3.5 ft/ $W^{1/3}$, a point was reached at which there was very small permanent displacement of the surface. In Figures 11 and 12 apparent crater dimensions of the Neptune 115-ton nuclear cratering shot in tuff [Shelton, Nordyke, and Goeckermann, 1960] have been included for comparison.

Early displacement of the surface as a function of time was obtained from high-speed motion-picture photography. Vertical displacements of the center of the mound immediately over the charge were measured. Velocities were obtained from the displacement-time curves. At early times surface vertical velocities are relatively constant. At later times, as the contribution of the gas bubble becomes evident, velocities in-

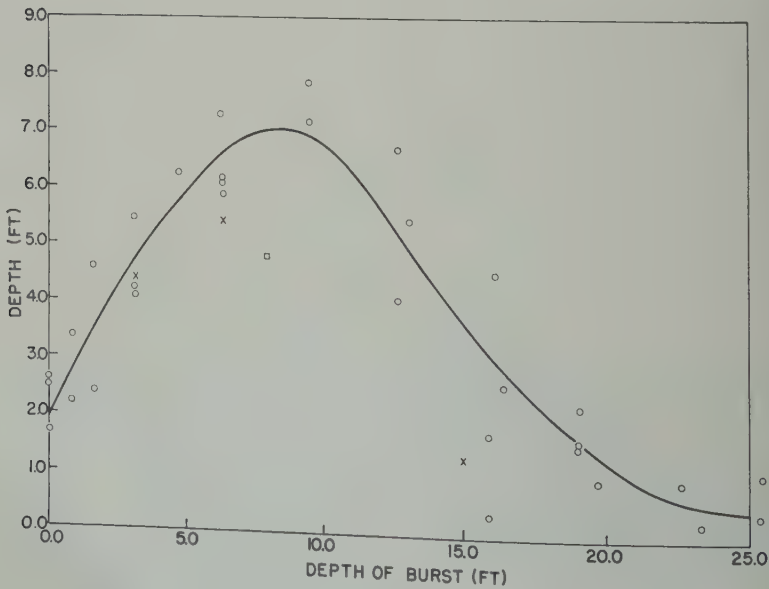
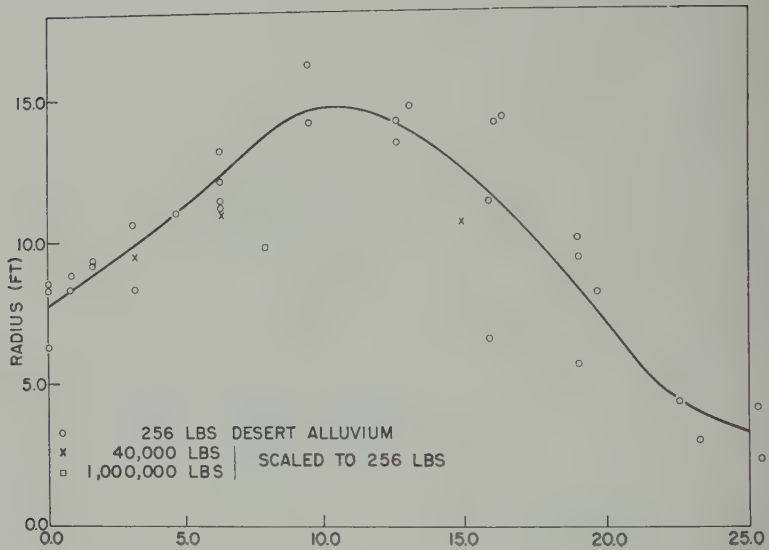


Fig. 5. Depth of burst curves, cube-root scaling.

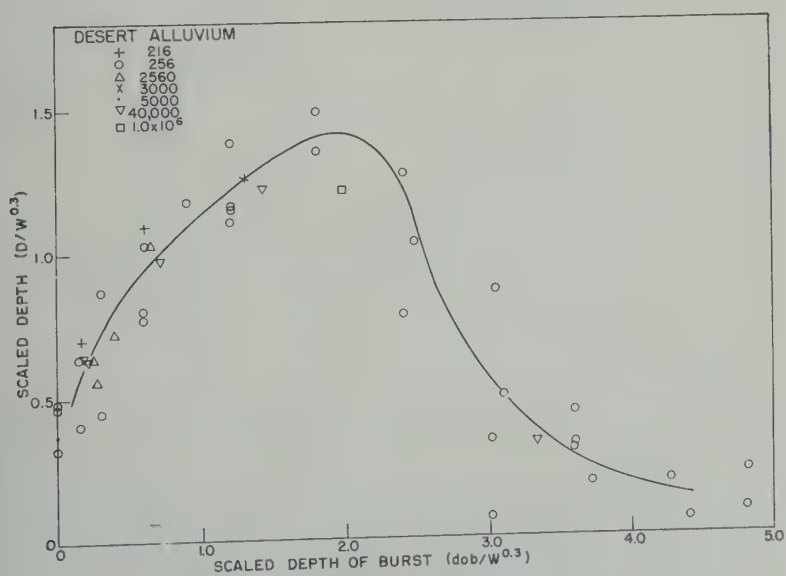
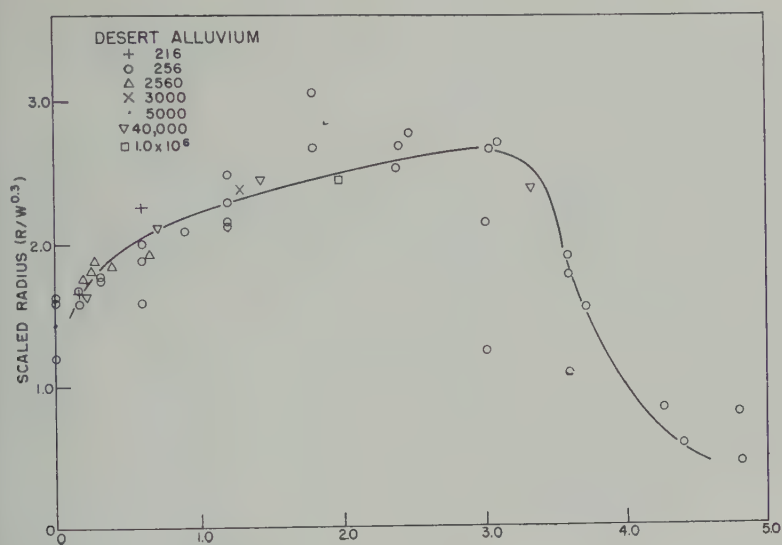


Fig. 6. Depth of burst curves, 0.3 scaling.

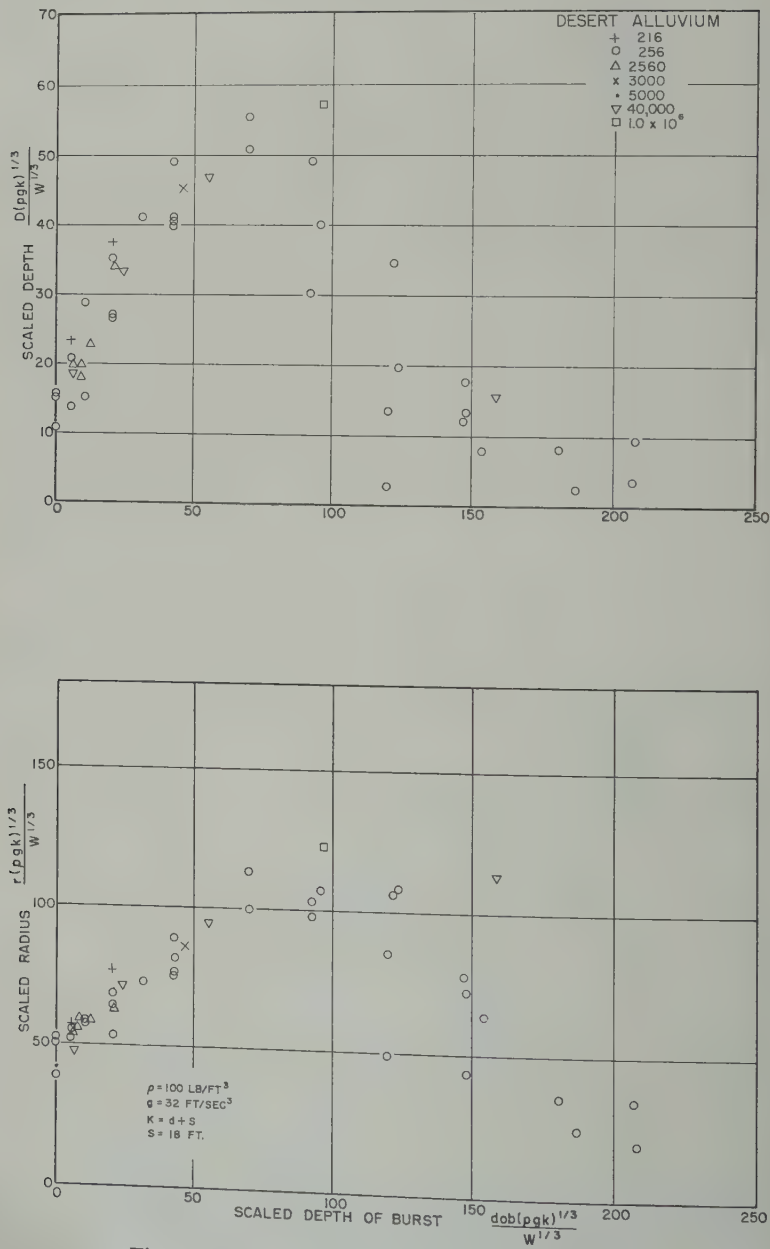


Fig. 7. Depth of burst curves, gravity considered.

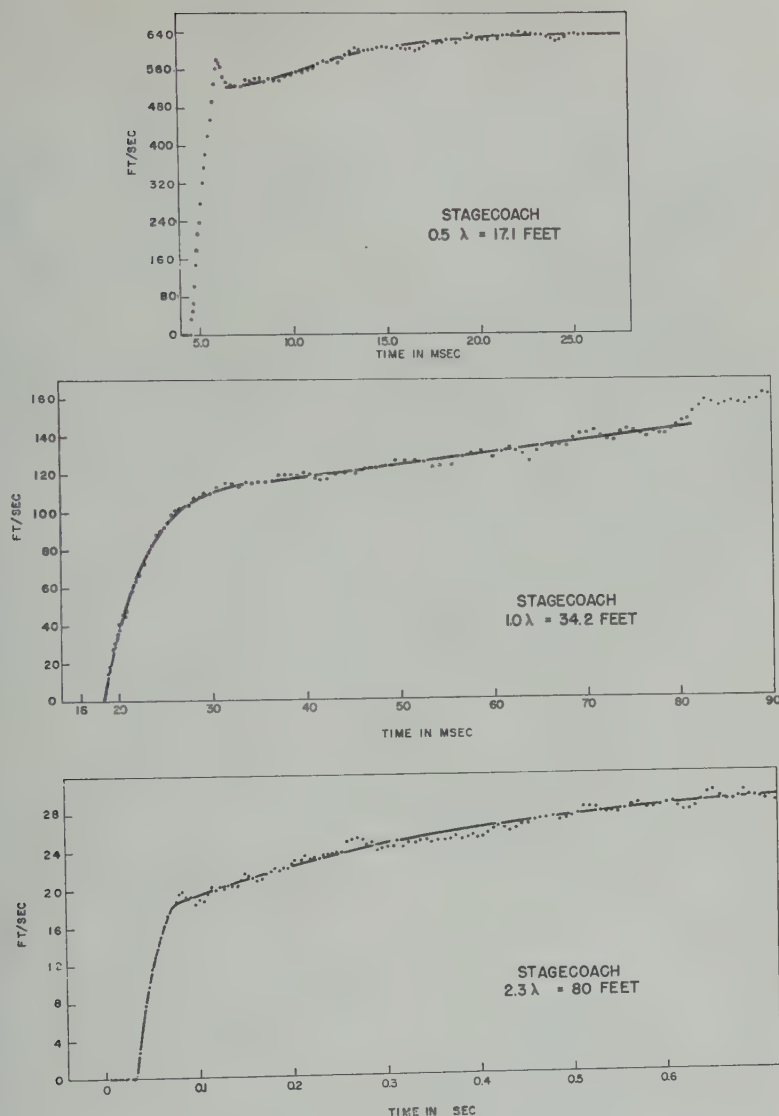


Fig. 8. Initial vertical surface velocities.

over those at early times. Typical values of early-time velocities are 80 ft/sec for a scaled burst depth of $1.5 \text{ ft}/W^{1/3}$ and 20 ft/sec for one at $2 \text{ ft}/W^{1/3}$. Figure 13 shows early velocities as functions of burst depth for 250-pound charges. Velocities are nearly the same as those in desert alluvium. Surface velocity for the Neptune nuclear shot has been included for comparison. The fact that surface velocities for the Neptune shot were

not greater than those for 256-pound high-explosive shots at comparable burst depths, even though Neptune crater dimensions were larger than those from the high-explosive shots at the same scaled burst depths, suggests that the slope under which the Neptune shot was fired may be one factor explaining why the Neptune crater was larger than expected.

Dimensions of the void below the surface were measured on one shot buried at $3 \text{ ft}/W^{1/3}$ scaled burst depth. A colored grout was pumped into

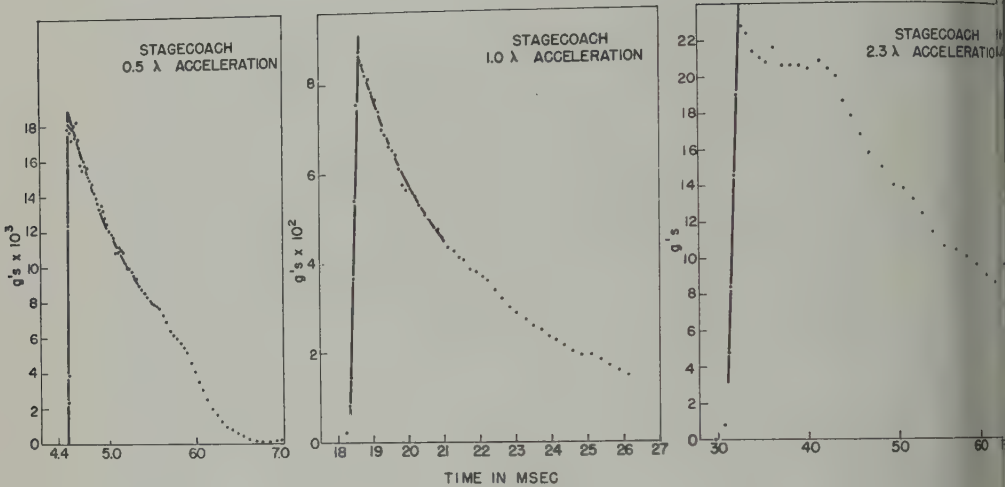


Fig. 9. Initial vertical surface accelerations.

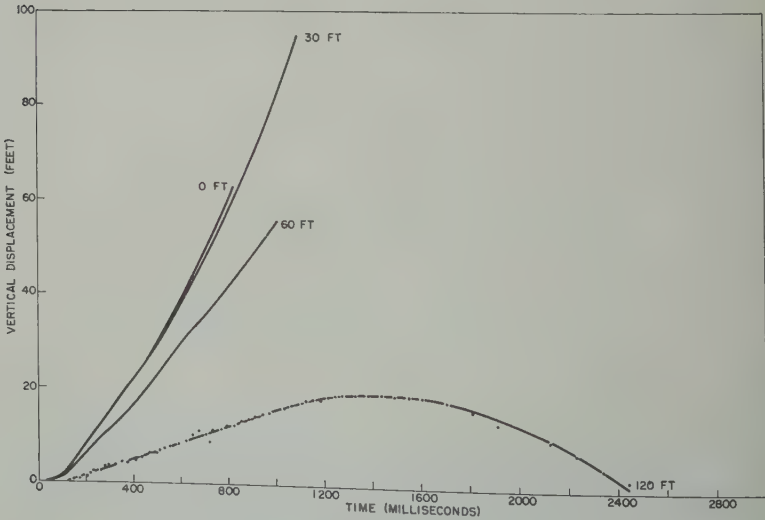
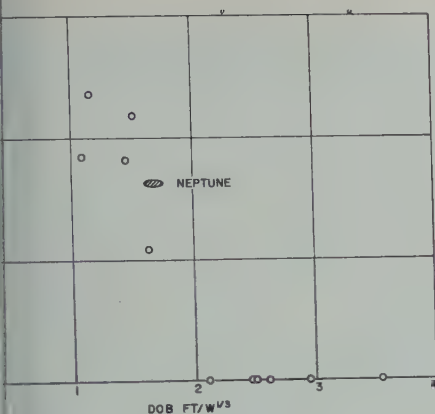
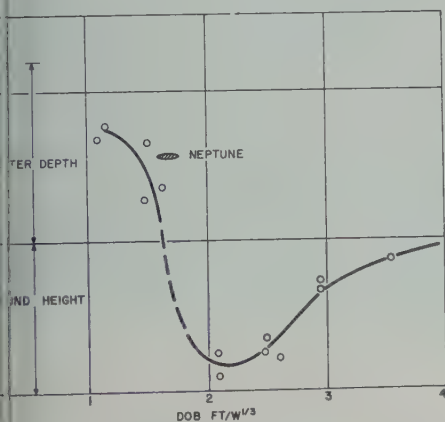


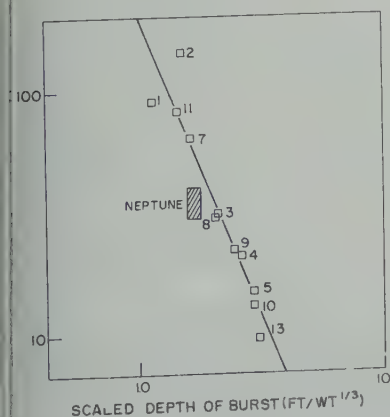
Fig. 10. Scooter target displacements.



Scaled apparent crater radius vs. scaled depth of burst.



Scaled apparent crater depth as a function of scaled depth of burst.



Early vertical velocity at surface zero vs. scaled depth of burst.

the void, and ten holes were core-drilled into the rock and grout. The shape of the void was determined from examination of the core. The volume of grout indicated a void of 263 cubic feet, approximating that of a spherical cavity having a 4-foot radius. Mapping of the grouted area indicated a volume comparable to that of a spherical cavity having a radius of 4.6 feet.

Typical sizes of the broken rock are a non-scaling feature and are in part predetermined by natural cracks and fissures in the medium. In a single material the distribution of rocks in each size range tends to remain relatively constant. This is especially true for the series of shots described here, all of which were the same size and placed at relatively deeper burst depths. Thus, larger charges may have given craters at depths deeper than $1.75 \text{ ft}/W^{1/3}$, since, because of size, the relatively smaller broken rock would not bridge or arch over the cavity and an apparent crater would be formed. This is a second factor that may account for the larger-than-expected Neptune crater.

We have concluded that larger charges are required to create craters that could be scaled with confidence to nuclear craters in tuff.

Dependence on basalt. Two series of high-explosive crater experiments have been fired in volcanic basalt. The first series was done in 1948 by the Corps of Engineers [The Panama Canal, 1948]. They used granular TNT in charges weighing from 8 to 200 pounds. The charge was cylindrical, its length was twice its diameter, and it was placed with the long axis vertical. Because of the charge shape, results from the Corps of Engineers experiments are utilized here primarily for comparison.

The second series was Project Buckboard [Vortman, Chabai, Perret, and Reed, 1960]. These shots were fired to determine crater dimensions in hard rock as a function of burst depth and to look for departures from cube-root scaling. The program consisted of ten 1000-pound charges, placed two each, 5, 10, 15, 20, and 25 feet deep, and three 40,000-pound charges placed at scaled depths of $\frac{3}{4}$, $1\frac{1}{4}$, and $1\frac{3}{4} \text{ ft}/W^{1/3}$. (Actual depths were 25.7, 42.8, and 60 feet.) One of the 1000-pound charges at 25 feet was a cylinder having a length-to-diameter ratio of 2. Its purpose was to provide a crater that could be compared with those of the earlier Isthmian Canal Studies experiment. All charges in Project

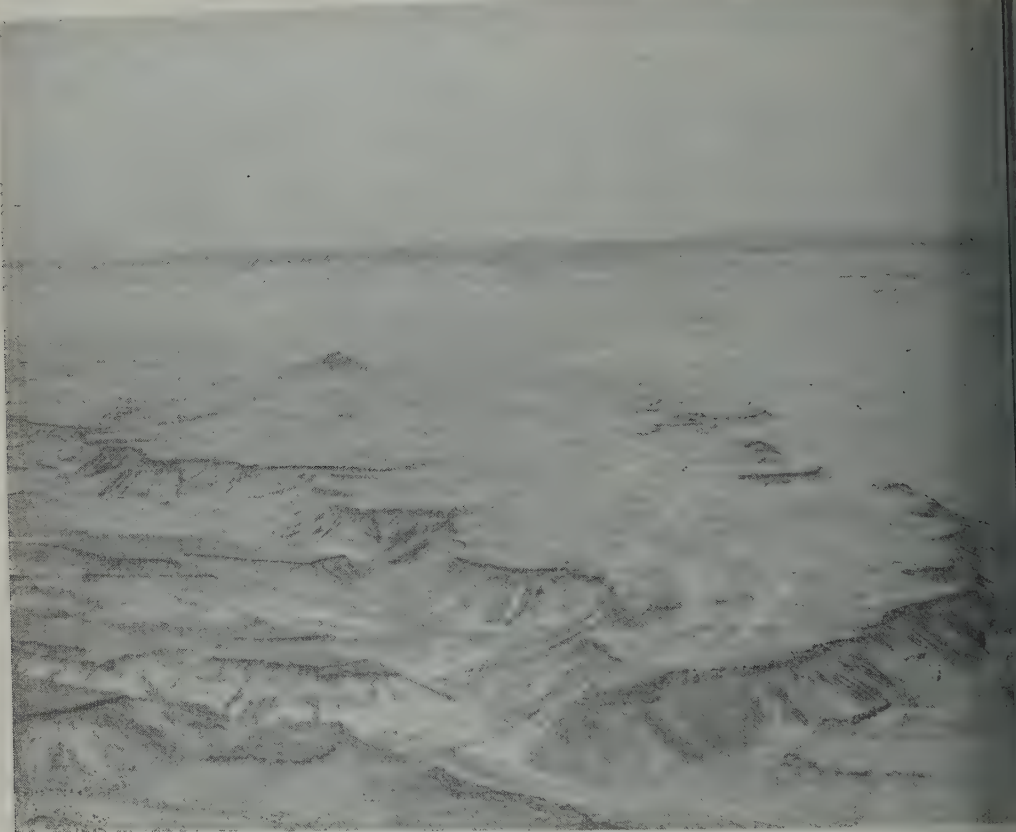


Fig. 14. Buckboard mesa.

Buckboard were covered with approximately 1 foot of dry sand over which the charge hole was stemmed with concrete.

The basalt was a Quaternary flow which had issued from a vent in the vicinity of a cinder cone near the north end of the flow and had flowed southward, presumably between two rows of low-lying hills which have since eroded. The remaining basalt-capped mesa (Fig. 14) rises 400 feet above the surrounding ground on the south end of the flow (foreground) and about 80 feet above it on the north end. In selecting the 13 firing sites, 37 exploratory NX holes were cored a total length of 1250 feet, the deepest to 125 feet. Shot locations were chosen from core tests where it was determined that the basalt near the depth of the charge had a density greater than 2.6 g/cm^3 and a compressive strength of approximately 20,000 psi.

Figures 15 and 16 show craters from the three larger shots. In Figure 15 the crater of the shallowest of the three 40,000-pound shots is shown at the greater distance. The crater of the deepest of the three is shown in the foreground. The striking difference between the two craters is the larger size of the ejecta in the case of the deeper shot. Rays of ejected material may also be seen surrounding the crater of the shallowest shot.

Figure 16 shows the crater from the shot of the middle depth, which was the largest of the three craters. There is some doubt whether the dimensions of this crater can be scaled with confidence, since it penetrated a pocket of cinder that had not been disclosed by preshot drilling. The crater was asymmetric. Its deepest point was located 10.6 feet from the vertical axis of the charge on a radius 9° north of east. The



Fig. 15. Photograph of Buckboard deep and shallow shots.

the surface contour was 3.5 feet from vertical axis and 10° west of north. Note the consequence of the cinder pocket the material ejected in the upper right of the photograph (south) are larger than the lower left.

apparent crater radius and depth as of scaled burst depth are shown in Fig. 7. Because of the scatter in data, it was not possible to detect any departures from cubing, and without such evidence we have no reason to continue its use.

With tuff, a nonscaling feature of the craters was size distribution of ejected material. With basalt, natural block-size distribution was partly predetermined by cooling in the basalt flow. At some of the deeper depths, the 1000-pound charges were not able to eject the largest blocks of

material; these fell or rolled back within the crater boundary, leaving a smaller apparent crater. If taken into account, the large blocks in the apparent crater greatly distort apparent crater dimensions. No rocks were found in the craters of 40,000-pound shots which, according to scale, were as large as those from 1000-pound shots, and those that were found produced no more than a small perturbation on the apparent crater. Thus, we have concluded that 1000-pound charges were not large enough to form a crater that could be scaled with confidence, especially where scaled burst depths were greater than $1 \text{ ft}/W^{1/3}$.

We did observe that, for a constant charge size, maximum rock size (and probably average rock size, also) increases with increased burst depth. Also, at the same scaled burst depth, maximum rock size increases as charge weight



Fig. 16. Photograph of Buckboard shot at 43 feet.

is increased. It should be clear that crater dimensions in basalt are smaller than those observed for desert alluvium.

We have also concluded that charge-shape effects on crater dimensions do exist. Crater dimensions from the single Buckboard cylindrical charge agree well with dimensions from those of the 1948 cylindrical-charge experiments [*The Panama Canal*, 1948] when cube-root scaling is used. Other types of scaling do not give as good agreement. We can conclude only that the cylinder gave a larger crater than the sphere of equal weight at the same burst depth ($2.5 \text{ ft/lb}^{1/3}$). However, there are reservations regarding the cylindrical-charge data from the 1948 experiment. Since in the Buckboard series we have concluded that 1000-pound charges were not sufficiently large for craters that could be scaled with confidence, the same statement must be

true to an even greater extent for craters from charge weights ranging from only 8 to 20 pounds. Only when a sound scaling method has been derived can an adequate comparison between the two experiments be made.

Diameter-to-depth ratio. The ratio of apparent crater radius to depth for craters in desert alluvium is shown in Figure 18. Similar ratios are obtained for craters in basalt, but no surface burst data are available. Note the scatter in values for surface bursts and the nearly constant ratio of 2 to 1 through the region of optimum crater dimensions.

Conclusions. Dependence of apparent crater dimensions in alluvium on energy releases over a range from 256 to 1,000,000 pounds can be described by the empirical rule that linear dimensions scale as $W^{0.3}$. A study of the effect of gravity in computing scaling indicates t

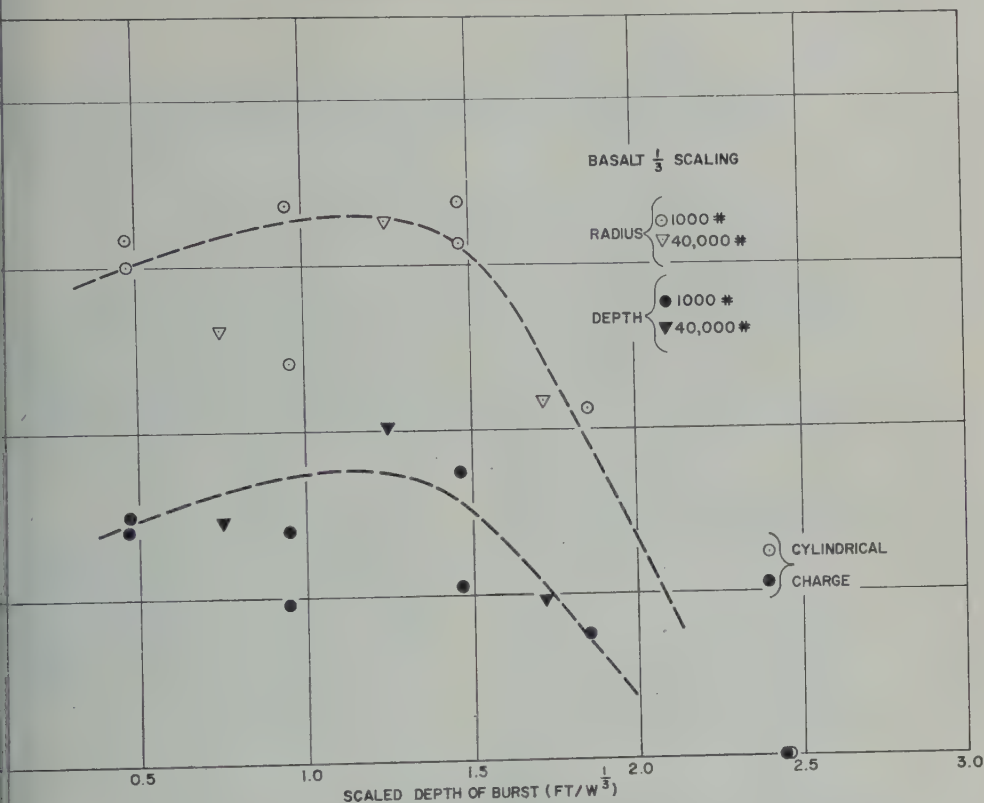


Fig. 17. Scaled apparent crater radius and depth as functions of scaled burst depth.

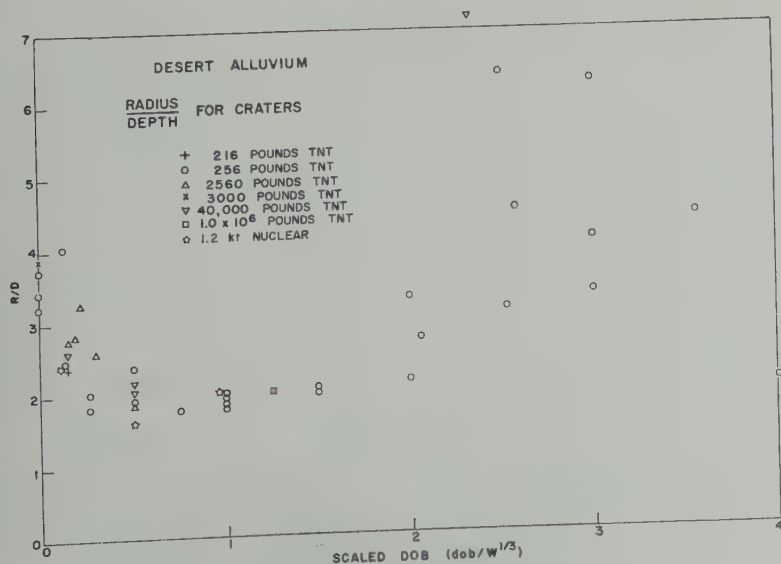


Fig. 18. Effect of depth of burst on radius-depth ratio.

reason for departure from cube-root scaling. Figure 6 provides a reasonable basis for comparison with nuclear explosives and delineates dependence of crater dimension on depth of burst. Comparison of Figures 6, 12, and 17 provides a basis for differentiation of crater dimensions in three different earth media. The natural deduction here is that, for rock materials, apparent crater dimensions should be determined from chemical explosions greater than 1000 pounds if the data are to be extrapolated to large yields.

Acknowledgments. The above research was performed under the auspices of the U. S. Atomic Energy Commission. The Scooter and Buckboard explosions were part of the Flowshare experimental program. Among the many who assisted in collection of these data, H. R. MacDougall, Sandia Corporation, who directed the field work, was particularly helpful.

REFERENCES

- Chabai, A. J., Crater scaling laws for desert alluvium, *Sandia Corp. SC-4391(RR)*, December 1959.
- Chabai, A. J., and D. M. Hankins, Gravity scaling laws for explosion craters, *Sandia Corp. SC-4541(RR)*, December 1960.
- Chabai, A. J., and L. J. Vortman, Project Stagecoach (interim report), *Sandia Corp. SCTM 181-60(51)*, May 1960.
- Feigenbaum, S. A., and P. L. Wegkamp, Photographic earth motion study, Scooter event, Edgerton, Germeshausen and Grier, L-510, February 1961.
- Johnson, G. W., Excavation with nuclear explosives, *Lawrence Radiation Lab. UCRL-5917*, refs. 1 and 8, November 1960.
- Murphey, B. F., High-explosive crater studies, *Sandia Corp. SC-4574(RR)*, May 1961.
- Murphey, B. F., and H. R. MacDougall, Crater studies: desert alluvium, *Sandia Corp. SCTM 119-59(51)*, May 1959.
- Perret, W. R., L. J. Vortman, A. J. Chabai, and J. W. Reed, Mechanisms of crater formation, Project Scooter, *Sandia Corp. SC-4602(RR)*, to be published.
- Sachs, D. C., and L. M. Swift, Small explosion tests, Project Mole, *Stanford Research Inst. AFSWP-291*, vol. II, December 1955.
- Shelton, A. V., M. D. Nordyke, and R. H. Goeckermann, The Neptune event—nuclear explosive cratering experiment, *Lawrence Radiation Lab. UCRL-5766*, April 19, 1960.
- The Panama Canal, Crater tests in basalt, *Isthmian Canal Studies, ICS Memo 284-P*, April 26, 1944.
- Vortman, L. J., A. J. Chabai, W. R. Perret, and J. W. Reed, Project Buckboard, interim report, *Sandia Corp. SC-4486(RR)*, November 1960.
- Warner, S. E., and C. E. Violet, Properties of the environment of underground nuclear detonation at the Nevada Test Site, *Lawrence Radiation Lab. UCRL-5542(Rev)*, April 1959.
- Waterways Experiment Station, Cratering from high explosive charges, compendium of crater data, *U. S. Army Engineers Tech. Rept. 2-54*, May 1960.

(Manuscript received June 5, 1961.)

Throwout Calculations for Explosion Craters¹

WILMOT N. HESS² AND M. D. NORDYKE

*Lawrence Radiation Laboratory, University of California
Livermore, California*

Abstract. This paper presents a study of the throwout from an explosion crater. Starting with particles in motion in the crater, the lip build-up is followed until all particles have landed. Comparing with experimental information from a 500-ton high-explosive explosion, we have attempted to calculate the apparent crater and lip shape and characteristics. So far the calculations are only rough, and the data are preliminary. We have changed g and also the size of the crater to see how craters on the moon would look.

To achieve a more detailed understanding of the cratering process, we have attempted to calculate the throwout from a crater and the subsequent build-up of the crater lip. Our approach is simple. We consider the material in the crater as being divided into zones and assign mass and velocity to each zone. We follow the particles through their parabolic motion until they hit the ground, and then we consider the construction of the lip.

Experimental information. Considerable experimental information is available to help in the calculation; one example is the data from the 500-ton Scooter TNT explosion carried out last year at the AEC test site in Nevada. The measurements made in connection with this explosion are of direct use to us here.

The motion of several targets near ground zero was recorded photographically. The photographs provide us with surface velocities at various positions, until gas venting obscures the data at about 1.2 seconds after the explosion (Figure 1).

From the photographs made at late time, we can follow individual rocks in the air and determine their velocities shortly before they hit the ground.

From the photographs we can determine a rough estimate of the maximum velocity of ejecta by measuring how high the cloud goes at early times (thermal effects are important).

Presented at the Geophysical Laboratory-Livermore Radiation Laboratory Cratering Symposium, held at the Geophysical Laboratory in Washington, D. C., on March 28 and 29, 1961. Also presented at NASA, Greenbelt, Md.

4. Twenty radioactive pellets were buried in the earth at known positions near ground zero before the explosion. After the explosion they were located by means of a counter and their final positions were measured. This information is shown in Figure 2, along with a profile of the apparent crater.

5. The final lip height was measured. Part of the increase in height close to the crater is caused by the upthrust of rock resulting from the explosion, and part is caused by the throwout of material.

6. One particular feature of the throwout, apparent from the photographs, agrees with information on terrestrial craters. The strata in the crater body appear inverted in the crater lip, at least close to the crater. Therefore material thrown out of the crater from the top of the crater landed first, and material from beneath it landed later. With all this information, we can learn a good deal about the throwout process.

Scooter calculation. Our objective in this study is to reproduce the Scooter throwout in detail. If we are successful in this, we should be able to calculate the throwout from other cratering explosions.

Our results of the Scooter calculation are preliminary. So far, we have considered that all particles are given an initial velocity and then undergo free fall, under the action of gravity alone. This 'impulsive' model we know to be wrong, as is shown by the early velocity history of one of the targets from Scooter in Figure 3. After the initial shock acceleration, the target slows down with $1g$ deceleration due to gravity until, at about 0.35 sec, a second acceleration

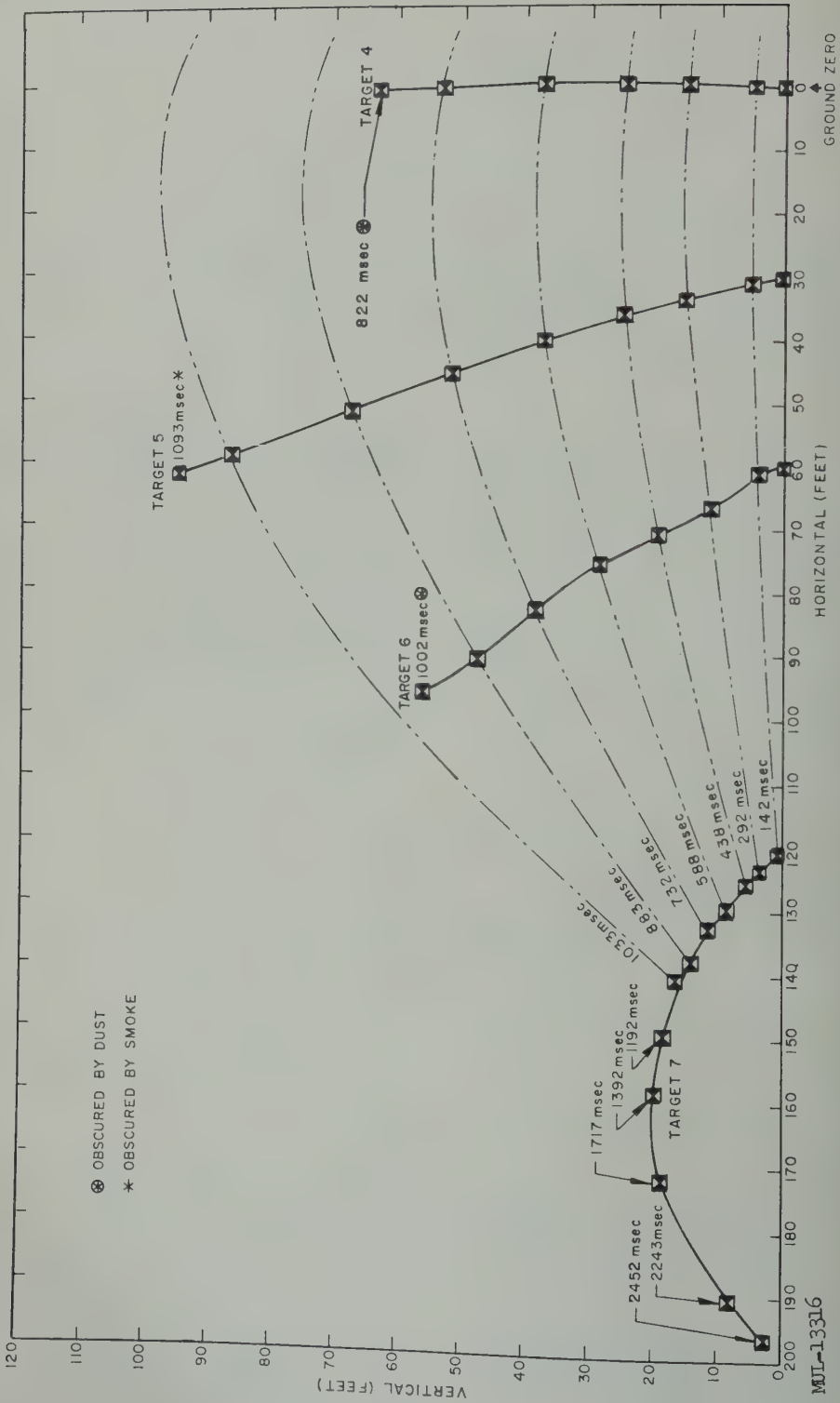


Fig. 1. Displacements of targets on the ground surface above the Scooter explosion vs. time.

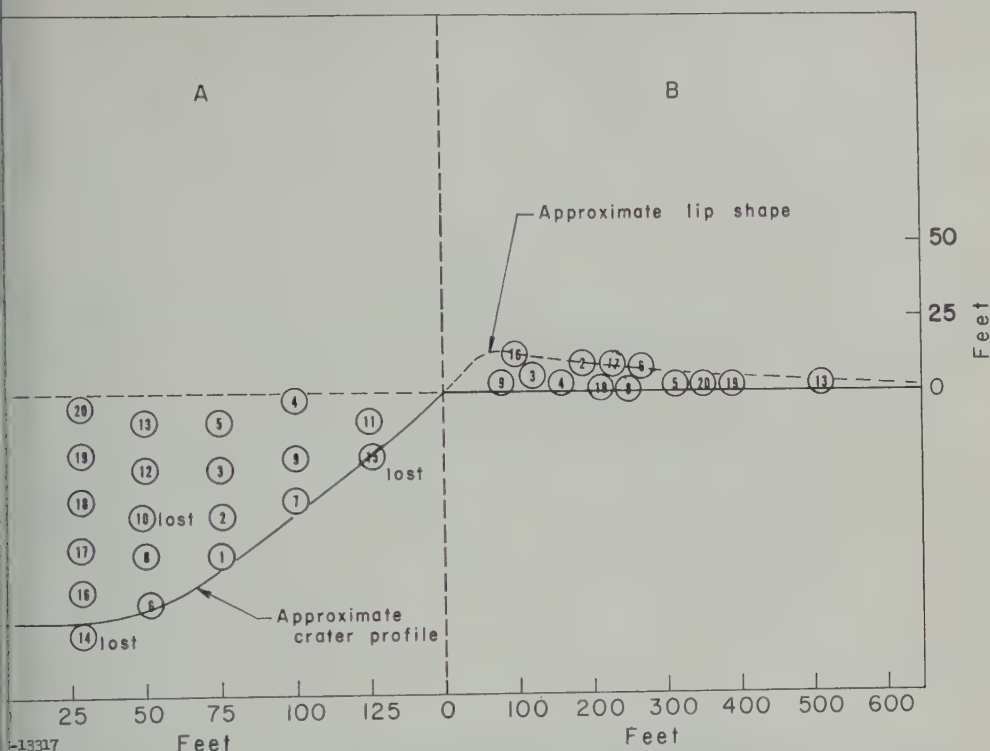


Fig. 2. Positions of the radioactive pellets before and after the Scooter explosion, showing (A) the pellets in place in the ground before the explosion, and (B) the approximate final positions of the pellets, with a different length scale. Pellets 1, 7, 11, and 12 remained in the crater and traveled only short distances.

by gas venting) starts and continues the record is obscured. We have also included aerodynamic drag and wind effects. These effects can and will be included in calculations. Our machine code can treat

a velocity that changes continuously with time. We have also found it necessary to include a consideration of the stability of talus slopes. The weak cementation and beds of relatively strong caliche in the medium allow the true crater to stand on a much steeper slope than is possible for the loose fallback material. Therefore fallback landing on this slope will slide down and into the center region of the crater. To treat this problem we have limited all apparent crater slopes to less than 0.75, the observed maximum slope in the Scooter crater. This has been done

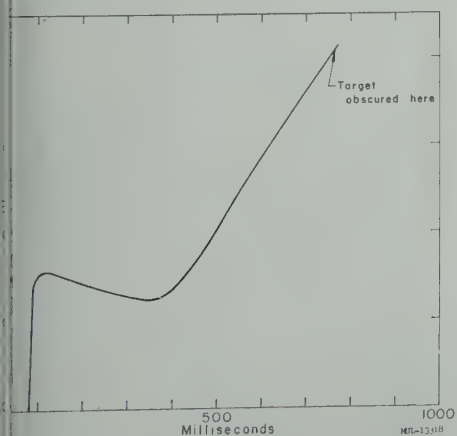


Fig. 3. Velocity history of a surface target at zero from the 500-ton Scooter TNT explosion.

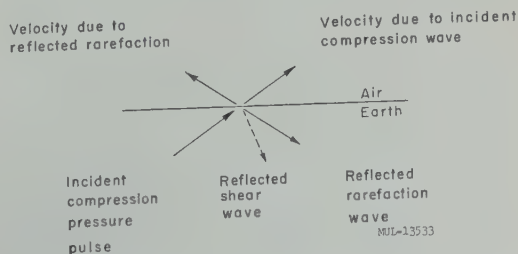


Fig. 4. Shock wave reflected at surface.

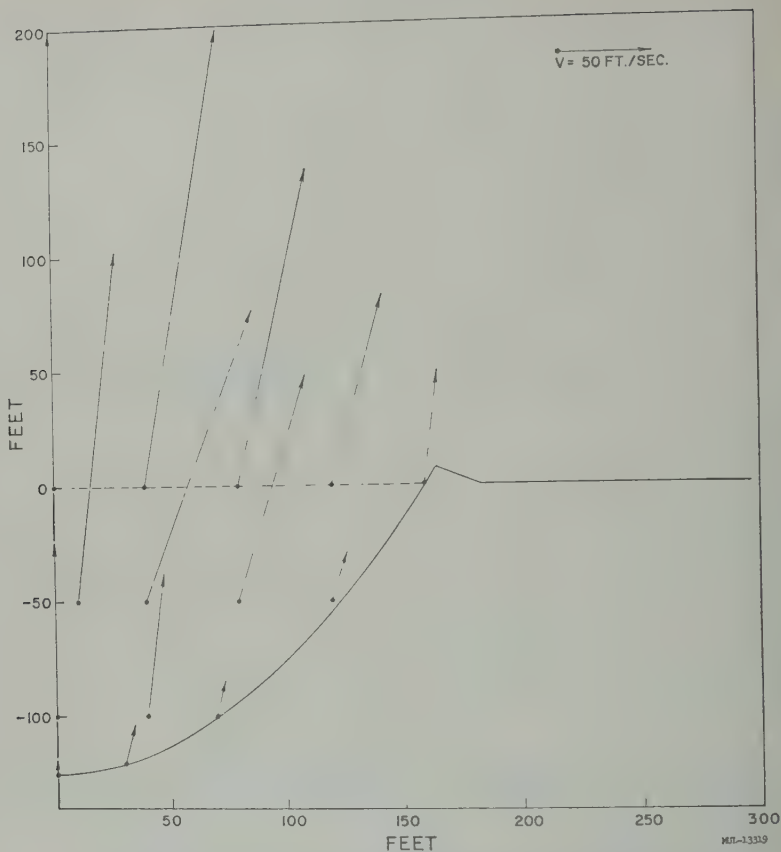


Fig. 5. The velocity field used in the throwout calculations. The ground-zero surface velocity is the largest at 200 ft/sec. The other velocities are to scale.

by spreading all excess material at a radius R uniformly over all the area inside of R .

In selecting the initial velocities in our impulsive model, we use experimental data from items 1 and 3 above. We have used a ground-zero surface velocity of 200 ft/sec, which gives a cloud height of about 600 feet, about the observed value. The velocities used decreased monotonically with radial distance from ground zero and also decreased with depth.

For positions away from ground zero, we must decide what direction to use for the velocity vector. We might instinctively make the velocity vectors radial outward from the explosion point. After a little thought, however, we might think that particles would leave the surface vertically because of the shock reflection conditions (see Fig. 4). Adding the velocity vectors due to the incident compression wave and reflected rarefaction should result in vertical motion of the

particles. But this still does not account for factors affecting the velocity vector. When reflection occurs, shear waves and surface waves are generated, which complicate the problem. This problem has been considered in the elastic case by *Cagniard* [1939]. Seidl (private communication, 1961) has solved Cagniard's equations and has found that the velocity vectors are nearly radial. Experimentally, the early velocities from Scooter are also radial, and the later velocities, when gas acceleration has occurred, are nearly radial as well.

In this, our first problem, we have used velocities that are more vertical than radial. For later problems we probably shall use radial velocities. The velocity field used for this problem is shown in Figure 5.

The calculation starts with all particles in the crater in motion, and the lip of the crater partially formed by upthrust (see Fig. 6, curve A). The

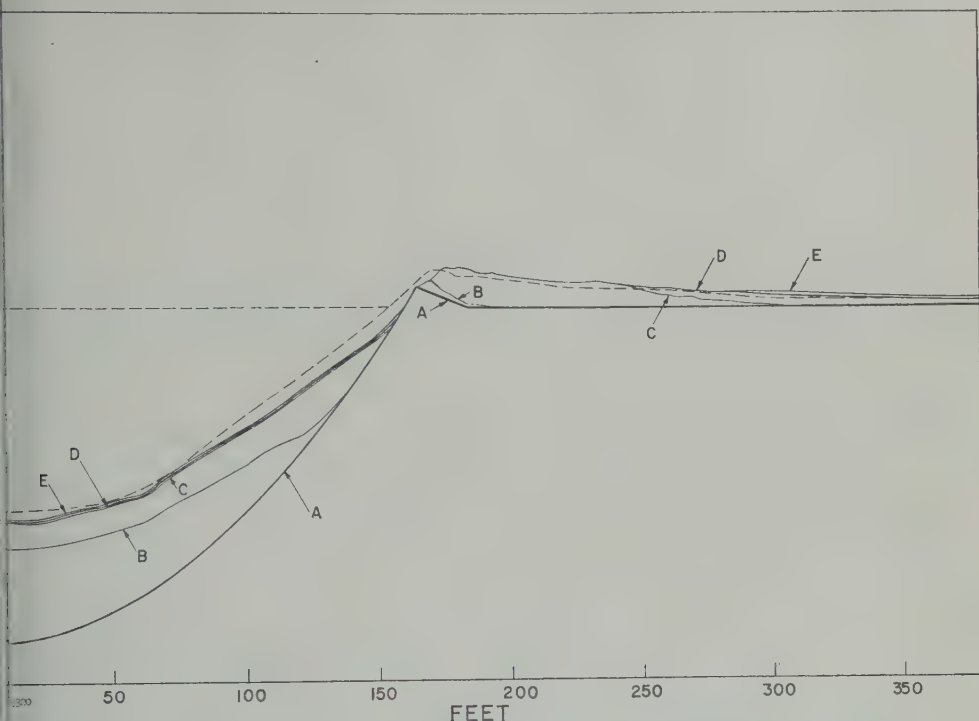


Fig. 6. The calculated Scooter crater profile as a function of time during the throwout period. Curve A is the true crater profile. Curves B, C, and D show the crater build-up at successively later times. Curve E is the final calculated crater shape.

of the velocity field is only roughly the crater of Scooter. We have assumed that below this line does not move any appreciable distance. We calculate where all the particles are after short steps in time, and, as they are hit in the crater or on the crater lip, we add out the mass to cover the proper range of positions and keep track of what particle is where, and on top of what other particle. The crater build-up are shown in Figure 6. Curve E we see the final apparent crater after all particles have landed. The dashed curve is the observed apparent crater profile for Scooter. Agreement here is quite encouraging.

Now that in several respects Figure 6 agrees with the actual Scooter throwout: (a) enough material falls on the outer edge of the crater, therefore some of the velocities are correct; (b) we do not get inverse stratigraphy, i.e., some of the velocities near the edge of the crater are ordered wrongly; and (c) the relative pellet positions from Scooter are not badly reproduced. All these discrepancies

can be corrected by properly modifying the velocity field in future problems.

Other calculations. We have run several more problems, starting with the same velocity field and varying other parameters. Even though we know that this velocity field is not wholly correct, it is not radically wrong, and scaling problems based on these velocities should be roughly correct. Figure 7 shows the apparent crater formed when the acceleration of gravity is reduced to 3.2 ft/sec^2 (curve C) compared with the case for $g = 32 \text{ ft/sec}^2$ (curve B). Most of the material gets out of the crater in C, and the apparent crater is nearly the same as the true crater.

We have also changed the scale of the problem in order to see how larger craters would look. In another problem, we have used 1600 feet as the diameter of the true crater while keeping the velocities the same as for Scooter. We think that, for the same scaled depth of burial, the velocities should stay constant for explosions of different sizes. Time and distance scale the same

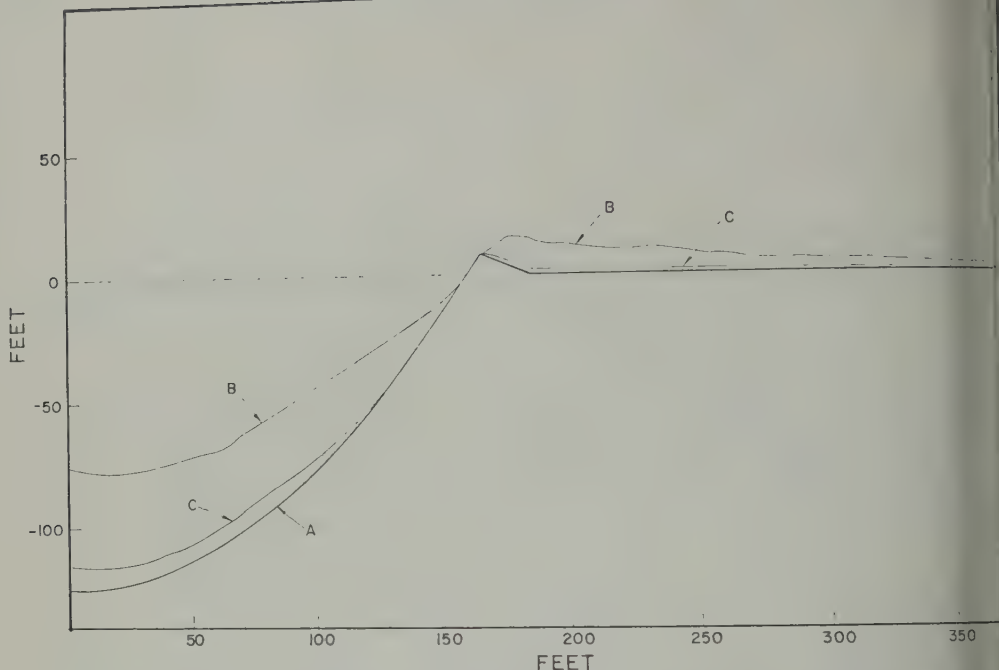


Fig. 7. A comparison of the calculated Scooter crater profile using normal earth gravity g (curve B) with the calculated profile using gravity of $g/10 = 3.2 \text{ ft/sec}^2$ (curve C). The true crater profile is shown in curve A.

way; therefore velocity, which is the ratio of these, should not change. The apparent crater from this explosion is compared with the Scooter calculation in Figure 8. The apparent crater is much smaller, as would be expected with particle velocities staying constant. The particles, therefore, move a smaller fraction of the crater radius, and hence more of them land within the crater. Changing the velocity field to get more mass on the outer edge of the crater, as suggested above, would eliminate the negative dip just inside the true crater lip. It can easily be seen that, if we had changed the scale another order of magnitude, we would have had no fallback on the upthrust lip of the true crater and a fairly flat region inside it. This would look like a walled plain on the moon.

Scaling. The dimensions of the true crater should vary as $W^{1/3}$, the cube root of the explosive yield. The true crater shape depends on the strength of materials, and it should show effects similar to the underground effects of explosions, which vary as $W^{1/3}$. Starting with this behavior for the true crater, we can get an

apparent crater scaling law from work like that shown in Figure 8. A calculation from Figure 8 gives an exponent of about $1/4$, somewhat smaller than expected. This number may well change when a new velocity field is used.

We can also consider the similarity of lunar and terrestrial craters. Starting with

$$h = v(ab \cos \eta)t - (g/2)t^2$$

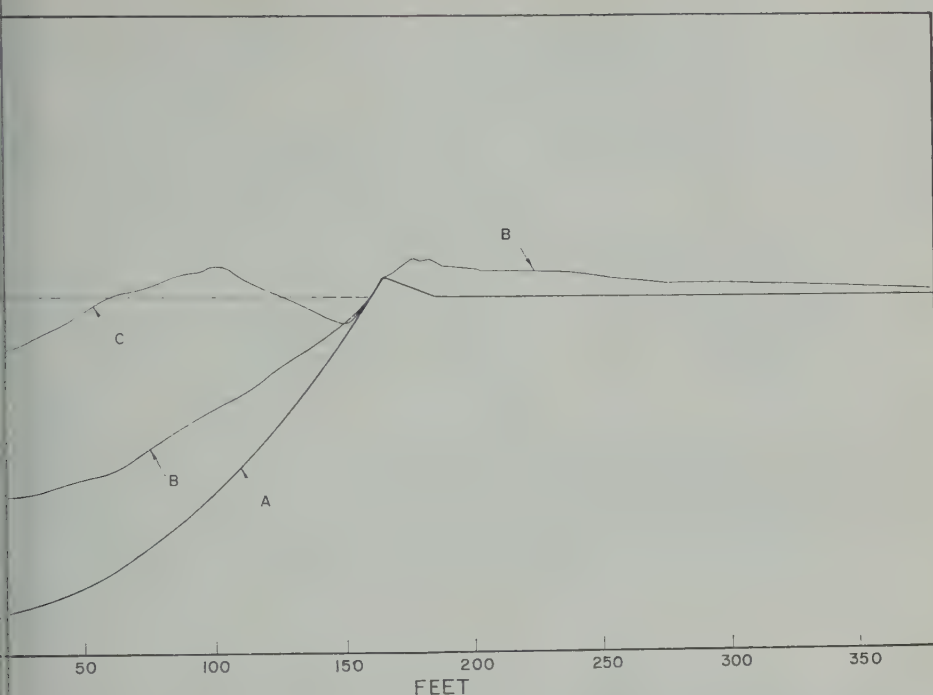
$$x = v(ab \sin \eta)t$$

where h and x are vertical and horizontal crater placements, v is the ground-zero surface velocity, a and b are functions that describe the change in velocity with radius and depth, η is the zenith angle of the velocity vector, we can get

$$gh = gx \cot \eta - g^2 x^2 \left(\frac{1}{2v^2 a^2 b^2 \sin^2 \eta} \right)$$

An interesting feature here is that dimensions x and h are both multiplied by g . For terrestrial and lunar apparent craters to be similar, we need

$$(l/R_{\text{true}})_1 = (l/R_{\text{true}})_2$$



Effect on apparent crater size of increasing size of true crater, calculated on the assumption that particle velocities are the same at comparable positions in the craters. Curve A represents the crater. Corresponding calculated apparent craters for Scooter (curve B) and for a true 600 ft in radius (curve C) show that the apparent crater becomes relatively smaller as the true size increases.

dimension l is either x or h . If this is satisfied, all the throwout particles are at the same fraction of the true radius. Then the total throwout distribution to the same true crater radius, will be the same for the two craters. We can achieve

$$(1\text{kt}, 1g) = \frac{2l}{2R_{\text{true}}} (8\text{kt}, \frac{1}{2}g)$$

has been doubled by increasing the field eightfold and l has been doubled. Because l and g appear as a product, decreasing one of them increases the other.

on the moon we get

$$(1\text{kt}, 1g) = \frac{6l}{6R_{\text{true}}} (216\text{kt}, g/6)$$

on the moon, where g is about $1/6$ that of earth, we could make a crater similar to one on earth by using 216 times the

explosive yield used on earth. The dimensions of the lunar crater would be 6 times those of the terrestrial crater, but the scaled sections of both craters would be identical.

One more point can be made about the relationship of terrestrial and lunar craters. For small vertical displacements ($h \approx 0$) we can write

$$x = \frac{2v^2}{g} a^2 b^2 \sin \eta \cos \eta$$

If on the moon there is no second acceleration due to gas venting (because no gas-producing materials such as water or carbonates are present), we might take the initial velocity of Scooter for explosions on the moon and the final velocity of Scooter for explosions on the earth. We get from this

$$\frac{x_{\text{earth}}}{x_{\text{moon}}} = \left(\frac{v_e}{v_m}\right)^2 \left(\frac{g_m}{g_e}\right) = \left(\frac{v_{\text{final}}}{v_{\text{initial}}}\right)^2 \left(\frac{1}{6}\right) \approx (2.5)^2 \left(\frac{1}{6}\right) \approx 1$$

This says that a crater on the moon from a certain explosive yield should be about the same size as a crater on the earth made by the same explosive yield. There may, of course, be some second acceleration on the moon which would make the lunar crater larger than the crater on earth.

Acknowledgment. This work was done at the auspices of the U. S. Atomic Energy Commission.

REFERENCE

Cagniard, L., *Réflexion et réfraction des ondes sismiques progressives*, Gauthier-Villars, 1939.

(Manuscript received July 6, 1961.)

Pacific Craters and Scaling Laws¹

R. B. VAILE, JR.

*Department of Physics, Stanford Research Institute
Menlo Park, California*

Crater measurements from two near-surface nuclear explosions detonated at Bikini 1954 were as follows:

Location	Approximate Yield, Mt	Crater Radius, ft	Estimated Maximum Depth, ft
Reef	15.	3000	240
Island	0.1	400	75

On the basis of these and additional crater data from previous nuclear detonations, an extrapolation procedure has been developed by which crater diameters can be predicted. This procedure is based on an empirical determination of the scaling exponent, m , as a function of soil type, using $W^{1/m}$, where R is radius, C is a constant related to the soil type, and W is the energy release. The range of uncertainty in the prediction of crater radius by this method is believed to be larger than a factor of 2.

INTRODUCTION

Current interest in craters revolves around a number of somewhat interrelated questions. The primary interest in craters is centered for the most part on the question of the size and shape of the crater produced by a specific military explosion detonated under specific circumstances. In the peaceful uses of explosives the question of cratering is of interest; in the military interest is the method of making the crater, but there are others, such as the question of minimizing fallout. From a scientific point of view, in order to deal with these questions in a comprehensive and satisfactory manner, it is necessary first to consider the detailed processes of crater formation and to establish the laws specifying crater shape and size. The primary interest in craters stems from the fact that protective structures, particularly those that are buried, are likely to survive an explosion with relatively little damage if they are only a short distance outside the zone of the realities of weapons delivery have attracted most of the military attention on craters from near-surface bursts. A major fraction

of the shots in the Pacific proving ground have been in this category, and this report covers in some detail those that have been declassified. Before reporting on the craters themselves it may be valuable to look into some aspects of crater nomenclature and theory.

In the investigation of craters formed by smaller explosions it has been recognized that, although the crater surface apparent to the eye is relatively easy to measure, there is nevertheless a disturbance in the earth, caused by the explosion, to some depth below this upper surface. The lower boundary of this volume of disturbed earth has become known as the 'true crater' in contradistinction to the upper surface, which has been called the 'apparent crater.' Although the term 'true crater' may be slightly misleading in its implications, it seems reasonably clear that for the purposes of determining the limitations of damage to underground fortifications the lower surface of the volume of disturbed earth (or true crater) is of greater significance than the apparent crater.

For small craters it is physically and economically feasible to determine the boundaries of both the apparent and the true craters, but for very large ones the excavations to determine the true crater would have to be too extensive to be practicable. The difficulties of measuring

¹Presented at the Geophysical Laboratory—Radiation Laboratory Cratering Symposium, held at the Geophysical Laboratory in Washington, D. C., on March 28 and 29, 1961.

the true crater become even greater if the crater is water-filled and if the level of radioactivity remains for some time high enough to prohibit extensive work.

The laws of similitude imply that the effects of an explosion of any (known) size in any medium are related precisely to the effects of an explosion of any other size in the same medium, provided that the medium fulfills certain rather stringent conditions. Experimental measurements using conventional explosives such as TNT lead to some optimism that craters produced by such explosives can be predicted with an accuracy almost entirely adequate for military purposes, even though it is clear that some properties of the medium (earth) in which the explosive is fired are very sensitive parameters in affecting the crater.

The situation regarding craters produced by nuclear explosives is less satisfactory. First, the evidence is meager, since, before Castle, there were only three such explosions on which crater measurements were made: Jangle U (underground); Jangle S (surface); and Ivy Mike. Second, the existing evidence leads to pessimism about the validity of scaling from conventional to nuclear explosion effects. The failure of crater scaling from conventional to nuclear explosions is believed to result from the enormous disparity in energy release (and this likewise applies between kiloton and megaton nuclear explosions) and also from the important difference in energy partition in the two types of explosions.

In general, the dimensions of the crater (radius or depth) are affected or determined by the total energy release, the depth of the charge, and the characteristics of the medium (earth) in which the charge is fired. If these parameters operate independently, an empirical equation could be written in the form

$$R = f(W) \cdot f(D_c) \cdot f(m)$$

or in the form

$$R = f(W) + f(D_c) + f(m)$$

where R is the radius; W is related to energy release, energy density, and detonation velocity; D_c is the depth of the charge; and m is related to the medium. In this case the separate contribution of each of the parameters can be determined easily. If, however, the parameters

are interdependent it is necessary to use the

$$R = f(W, D_c, m)$$

and the effect of varying any one of the parameters is much more complicated because it depends on the values at which the other parameters are maintained.

There is general agreement among investigators that the parameters affecting craters in fact extensively interrelated. The universal use of scaling concepts, particularly the scaling of depth of charge, is evidence in point. Thus, with respect to the effect of energy release and depth of charge a satisfactory form for the equation

$$R = f(W) \cdot f(W \cdot D_c)$$

or, as a more specific example,

$$R = W^{1/k} \cdot f(\lambda_c)$$

where k is approximately 3 and

$$\lambda_c = \frac{\text{Charge depth, ft}}{(\text{Weight of TNT, lb})^{1/3}}$$

The inclusion of an additional term to represent the effect of different mediums could be in several forms, among them

$$R = f(W) \cdot f(W, D_c) \cdot f(m)$$

$$R = f(W, m) \cdot f(W, D_c)$$

The data at hand have seemed to the author to fit better into an equation of form 2 than into one of form 1, namely,

$$R = (WE)^{1/m} \cdot f(\lambda_c)$$

as will be elaborated later. It is to be noted that these two forms are drastically different in their implications of extrapolation from less than kiloton charges up to megaton charges.

In attempting to correlate crater data from TNT blasts with those from nuclear explosions it has in the past appeared useful to include a factor less than unity (0.3 to 0.9) in the value of W assigned to nuclear charges in terms of equivalent tons of TNT, based on radiochemical data. This has been justified by the fact that the energy partition is totally different in the two types of explosives and that the nuclear weapons deliver radiant energy whereas conventional explosives do not. It is believed, however, that, at best, correlation will be uncertain.



Fig. 1. Castle 1, preshot photograph (Fig. 3.2, WT-920).

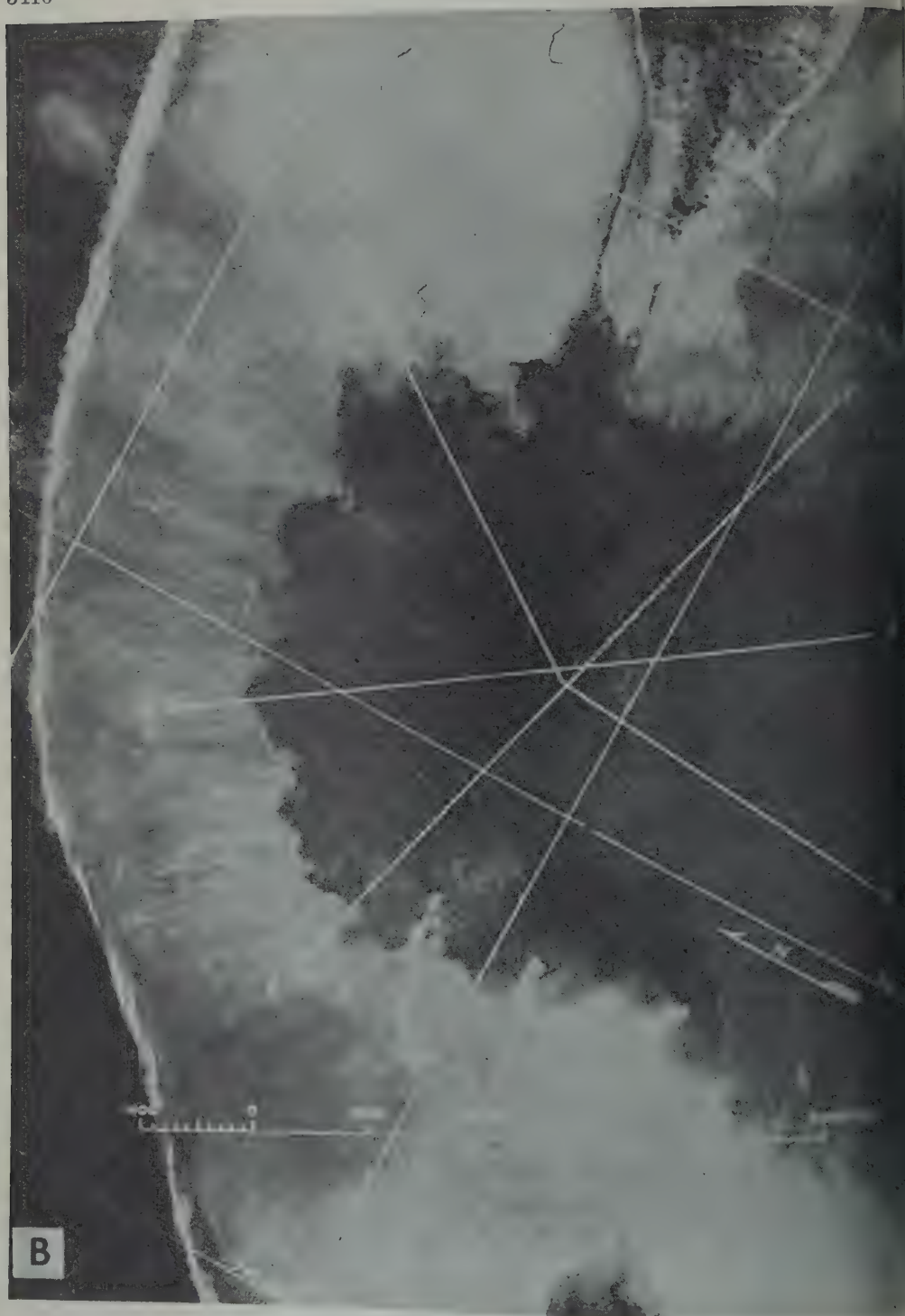


Fig. 2. Castle 1, postshot photograph (Fig. 3.3, WT-920).



Fig. 3. Castle 3, preshot photograph (Fig. 3.6, WT-920).

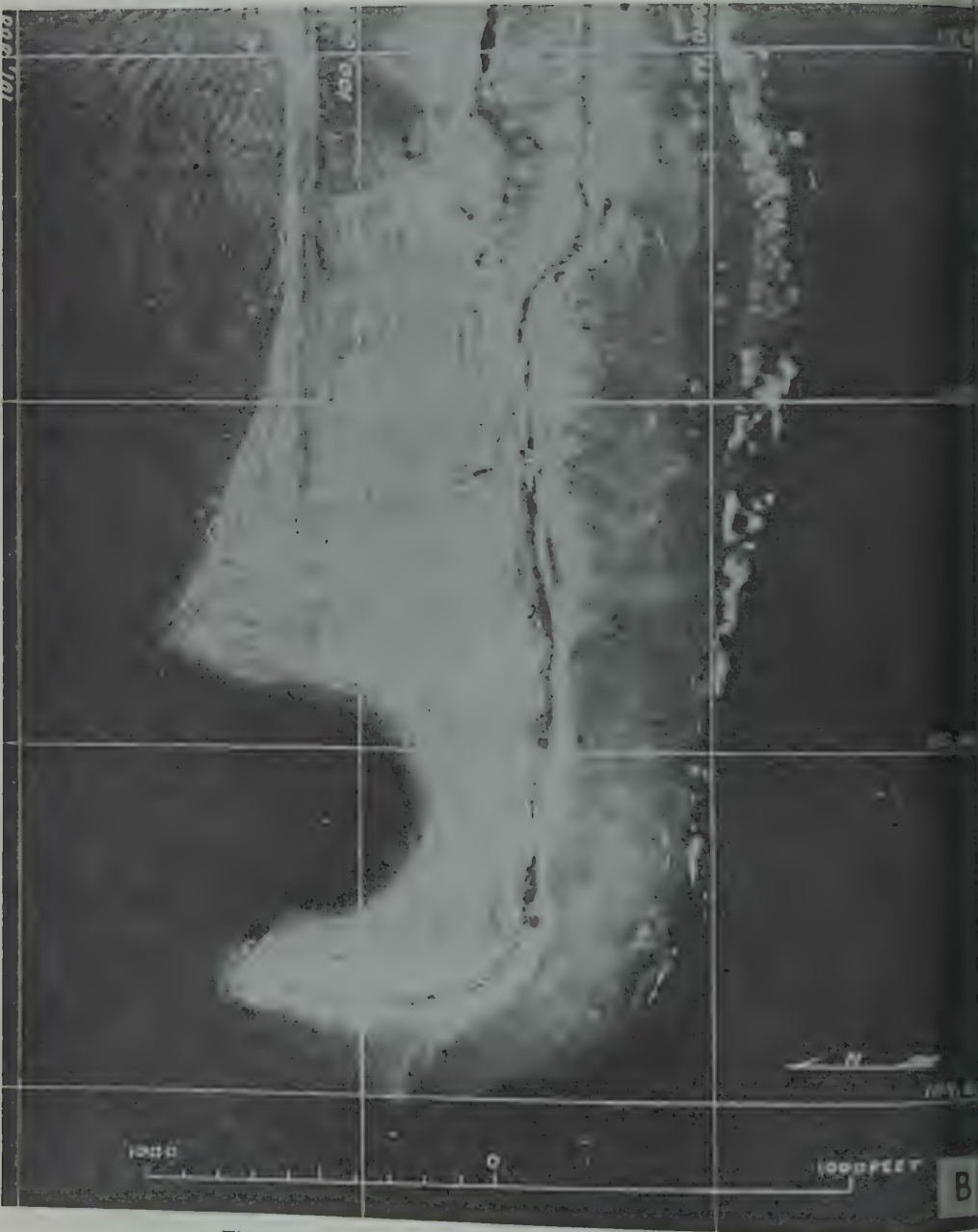


Fig. 4. Castle 3, postshot photograph (Fig. 3.7, WT-920).

the advent of megaton weapons the effect of sizes is so great that good correlation cannot be expected.

Effect of charge depth (or height) (λ_c) is well established for TNT. If scaled crater diameter is plotted against scaled charge depth, the curve is plotted against scaled charge depth, near both from experiment and from reasoning that the curve will be concave downward, since no surface crater is formed if the charge is sufficiently high above the surface or sufficiently deep below it. For the maximum of this curve is rather broad and occurs in the range of $1 < \lambda_c < 3$, where $\lambda_c = t/(1b \text{ TNT})^{1/3}$.

The effect of the medium, $f(m)$, has been shown to be as large as a factor of 2 in field experiments with TNT. Unfortunately, the specific properties of the medium which affect the crater are not yet established. It is postulated that strength, either shear or tension, and density are sensitive parameters. Possibly the elastic moduli are also important. In regard to strength, it is, of course, the strength under shock load conditions that is important. To make laboratory tests under shock load conditions is difficult, and the heterogeneous nature of earth makes the extrapolation from laboratory to field conditions very uncertain. Thus, although

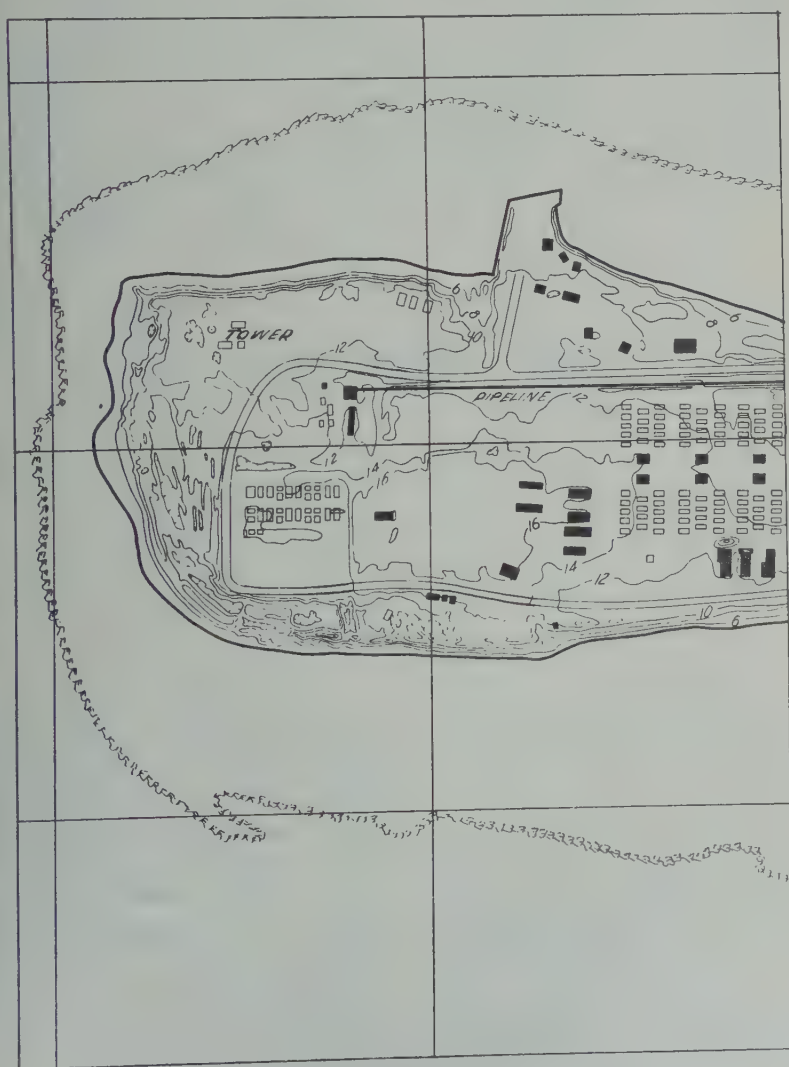


Fig. 5. Castle 3, preshot map (Fig. 3.8, WT-920).

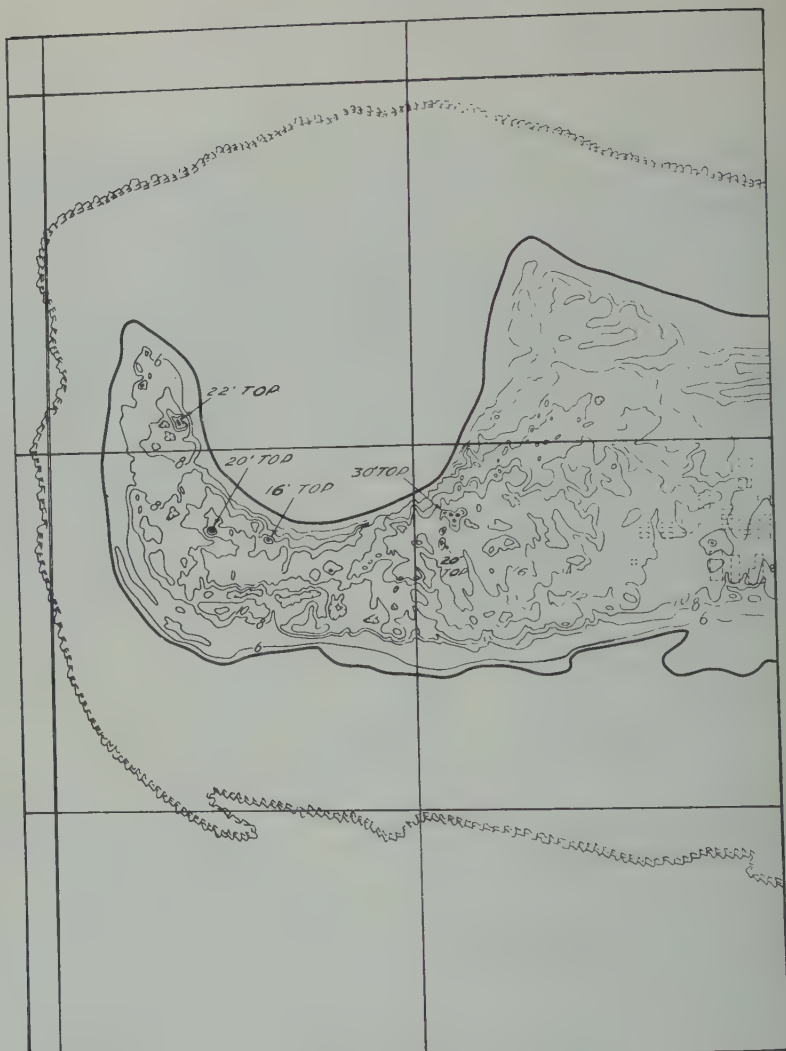


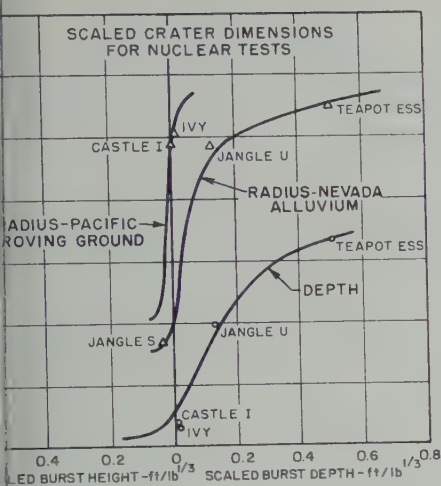
Fig. 6. Castle 3, postshot map (Fig. 3.9, WT-920).

appropriate values for strength under shock load are not known, clearly the strength under such conditions may differ markedly from the strength under static load.

In theory, the density of the medium may significantly affect crater size. In practice, however, the range of densities found is trivial compared with the range of strengths, and hence the density is believed to be a parameter of only minor importance in affecting the crater.

As has been mentioned, the application of similitude principles places certain requirements on the medium. At a minimum for the purposes

of crater investigation, the properties of the medium at equivalent locations (scaled) in two experiments must be identical. This requirement is completely met if the two media are homogeneous, isotropic, and identical. The properties of earth, however, are greatly affected by overburden pressure. Thus in a static sense the properties of earth are grossly dependent on actual (not scaled) depth below the surface, and in a dynamic sense these properties will be similarly affected by the pressure produced by the explosion. Thus one of the fundamental conditions for the proper application of similitude



Scaled radius and depth vs. shot depth.

laws is violated. The greater the range of explosion, and hence of depth, the more this violation becomes.

Further difficulty with the application of occurs in situations such as existed on where two media, earth and water, were l, and where the earth was saturated so rces were transmitted by a complicated ation of intergranular forces and hydraulic es.

ic craters. Most of the shots in the have been over water areas, and though ay have had some influence on the bottom s they did not produce craters in the tional sense. Others have been fired at gh altitudes that no physical surface

effects occurred. Still others have been fired above land areas but at a height such that the craters were of the depression type rather than the scoured or throwout type. There remain shots that have been fired on land areas and whose yields and results have been declassified. This is the group with which we are concerned here.

Measurement of craters at the Pacific proving ground is fraught with considerable difficulty inasmuch as all are water-filled and most of them have been washed by waves before it has been possible to measure them. The presence of water greatly increases the uncertainties in the actual dimensions of a crater. It also necessitates quite different techniques in measurement. The techniques in the Pacific are primarily either stereographic photography from the air or measurement of water depth inside the crater by means of a boat-carried fathometer. Incidentally, in this procedure there is greater difficulty in determining the horizontal position of the boat than in determining the depth of water under it.

The largest crater in the Pacific was produced by shot Castle 1, for which the preshot aerial view is Figure 1. This shot, which was approximately 15 megatons in yield, produced the crater indicated in Figure 2, which as can be seen is approximately 3000 feet in radius. At the time this crater was measured, approximately a week after the shot, the crater area had been washed to the extent that the fathometer showed a very flat bottom at 100-foot depth. Extrapolation of the slope of the sides where they could be

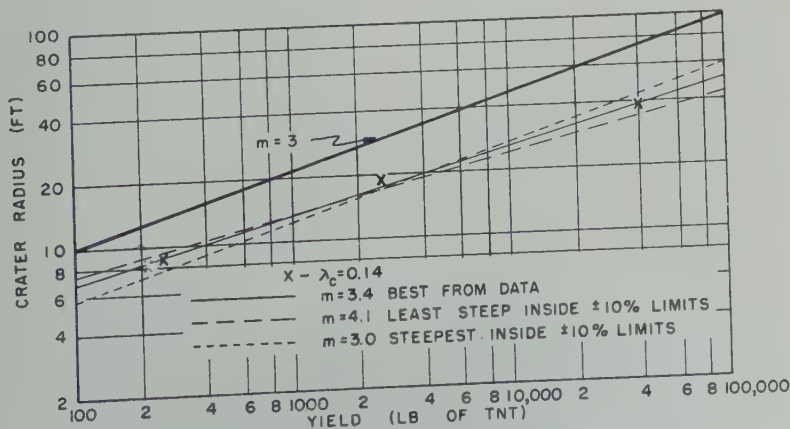


Fig. 8. Crater radius vs. yield, Nevada, $\lambda_c = 0.14$ (Fig. 4.1, WT-920).

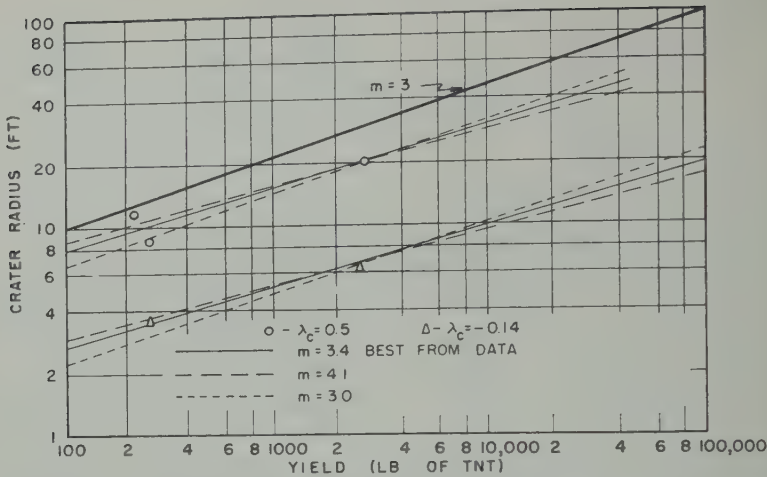


Fig. 9. Crater radius vs. yield, Nevada, $\lambda_c = 0.5$ and $\lambda_c = -0.14$ (Fig. 4.2, WT-920).

measured combined with information on the length of anchor chain required to moor a barge in the center of this crater for a following shot led to the conclusion that the real depth at the time of formation was about 240 feet.

A similar shot fired some years earlier in the Ivy Operation, Ivy Mike, produced a crater at about the same radius and of somewhat shallower depth. The shallower depth is undoubtedly due to the fact that the Ivy Mike shot was fired some 20 feet above the surface whereas the Castle 1 shot was much closer to the surface.

Castle shot 3 of about 100 kt was fired near the northern end of Tare Island on Bikini atoll as indicated in preshot aerial photograph, Figure 3. The result, a crater of 400-foot radius, some 75 feet deep, is shown in the aerial photograph of Figure 4. Maps made from these photographs and from pre- and postshot surveys are shown in Figures 5 and 6. This crater, like all the others, was water-washed by the wave resulting from the shot that produced the crater, but this one was washed in addition by the wave from a later shot fired on the surface of the lagoon, and so the crater outline had been blurred to an indefinite extent before the measurement could be made.

The great sensitivity of height or depth of burst, particularly in the region very close to the surface, is well illustrated in Figure 7, in which the upper curves are scaled crater radius plotted against scaled depth of charge and the lowest curve is scaled depth of crater against

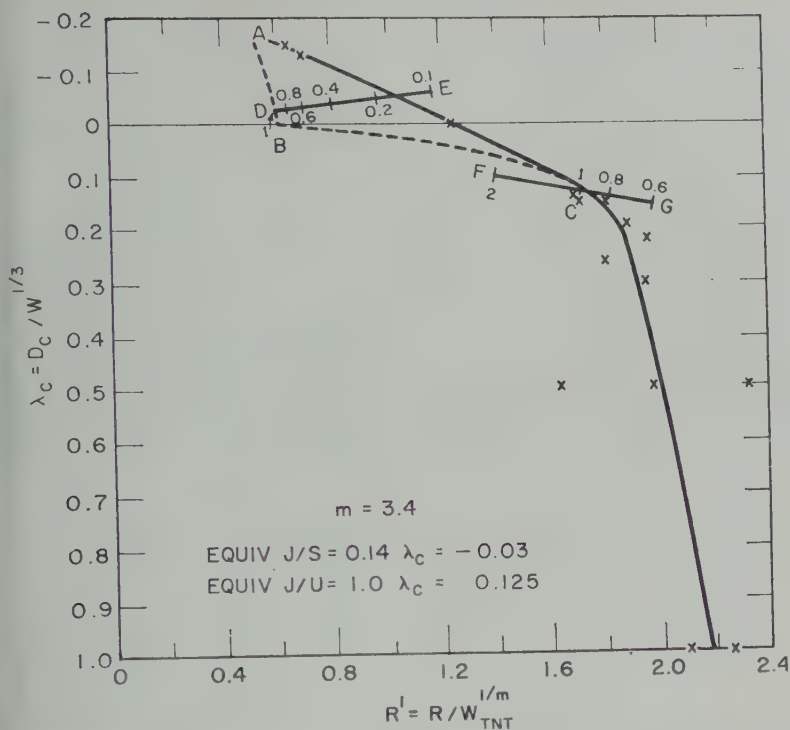
scaled depth of charge. The declassified nuclear tests in the Pacific including the ones illustrated above are shown here, and for comparison the nuclear cratering shots at the Nevada test site are also shown. The two elements of interest in this figure are the steep slope of the radius curves in the vicinity of zero depth and the fact that even with the steep slope the range of uncertainty totals no more than about a factor of 3.

PREDICTION OF CRATERS

The crater prediction method presented here typifies the procedures recommended and is none as precise as it might be if more recent crater data were included. The pressure of time in the original project precluded treatment of scaling crater depth; only crater radius is so treated.

Background. The data required about any specific explosion for which a prediction of the crater is desired are (1) the yield, (2) the type of soil, and (3) the depth or height of burst. With this information, it is appropriate to look at the existing evidence and measurements and to develop rational procedures for extrapolation or interpolation.

The craters from explosions high above the surface are significantly different from those formed by lower explosions in that they are depressions rather than excavations. Such craters being of relatively minor importance from military standpoint, are not considered here. Also, since an attempt to distinguish true from



Scaled crater radius vs. scaled charge depth, Nevada alluvium ($m=3.4$) (Fig. 4.3, WT-920).

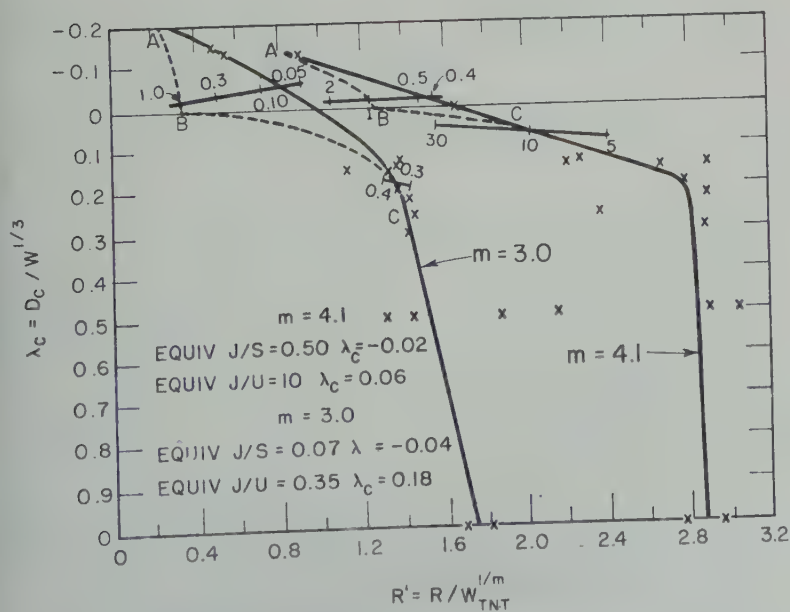


Fig. 11. Scaled crater radius vs. scaled charge depth, Nevada alluvium ($m=3.0, m=4.1$) (Fig. 4.4, WT-920).

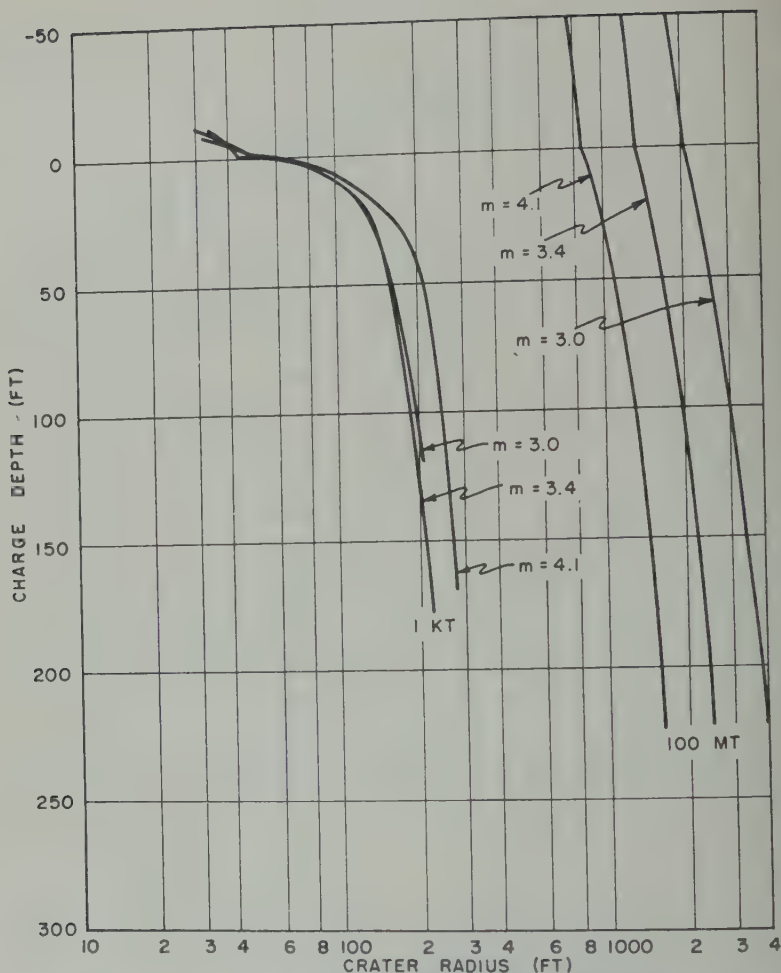


Fig. 12. Crater radius vs. charge depth, Nevada alluvium showing range of uncertainty (Fig. 4.5, WT-920).

apparent craters becomes less and less realistic as larger and larger yields are considered, only apparent craters are considered.

In previous analyses of crater data, the horizontal dimension used has sometimes been diameter and sometimes radius, and these values have been measured sometimes from lip to lip and sometimes at the original ground level. In this paper, only radius at original ground level is considered.

In reviewing the existing data from a broad point of view and with the objective of crater prediction for megaton explosions in mind, the following facts stand out:

1. All the data from which soils can be com-

pared are contained in experiments involving relatively small quantities of TNT.

2. When more than one explosion has been fired under presumably identical conditions, important scatter of the dimensions of the resulting craters is apparent.

3. The range over which these data must be extrapolated to permit prediction of megaton craters is enormously greater than the range of extrapolation commonly accomplished in engineering or scientific fields. The situation is roughly equivalent to an attempt to predict the penetration of the projectile from a new tank gun through armorplate on the basis of observation of many measurements of the

of BB's from an air rifle through tin
a few measurements of the penetration
of bullets through pine.

result of these facts any extrapolation
is inevitably associated with a large
uncertainty in the final result. In making any
prediction, consequently, it is of major
importance to indicate the order of magnitude
of uncertainty involved as well as the
prediction itself.

At the outset of any attempt to develop
prediction procedures, one is faced with a
critical choice. On the one hand he may
delve into the mechanism of the phe-
nomenon, and, on the basis of physical analysis,
predict the causes, the effects, and the influence

of specific parameters. Alternatively, he may
adopt the attitude that, in a complicated phe-
nomenon such as crater formation, the mechan-
isms by which causes and effects are interrelated
are so little known as to be, for the moment,
unknowable, and hence conclude that the
appropriate approach is the empirical extrapola-
tion of the existing data into the range of
parameters where prediction is desired. It is my
opinion that the second approach is the more
realistic one under the circumstances involved
in the present problem, and that is the approach
I have taken. The most important deviation
from past thinking occasioned by this approach
is that cube-root scaling is, on this basis, dis-
carded as a primary tool in the extrapolation

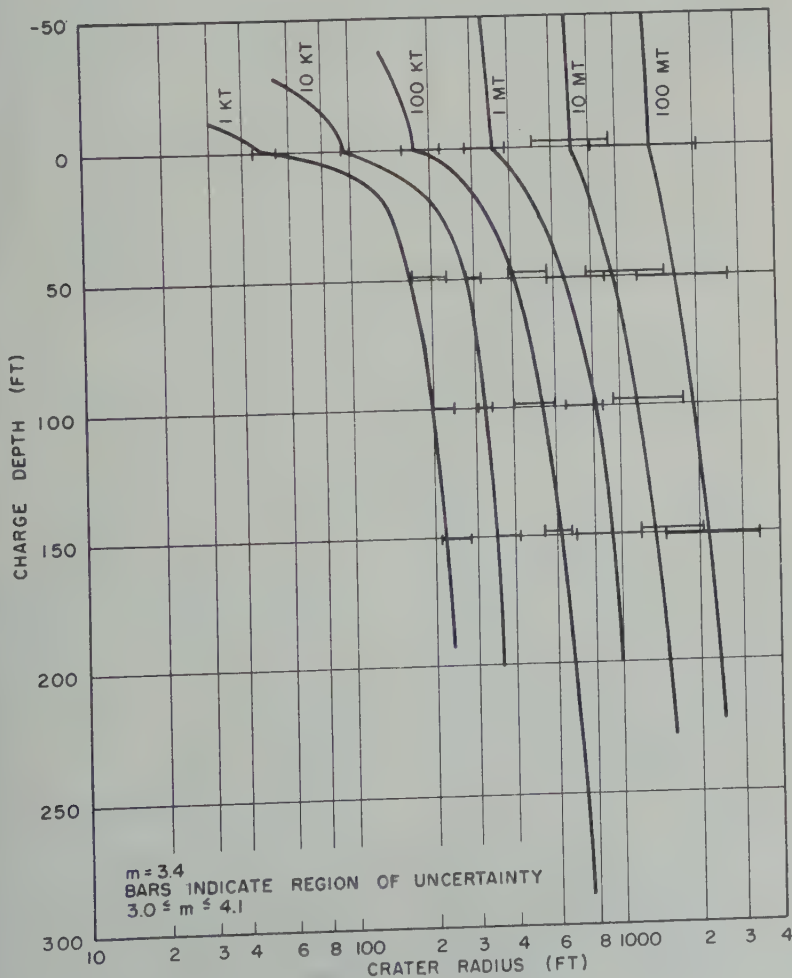


Fig. 13. Crater radius vs. charge depth, Nevada alluvium (Fig. 4.6, WT-920).

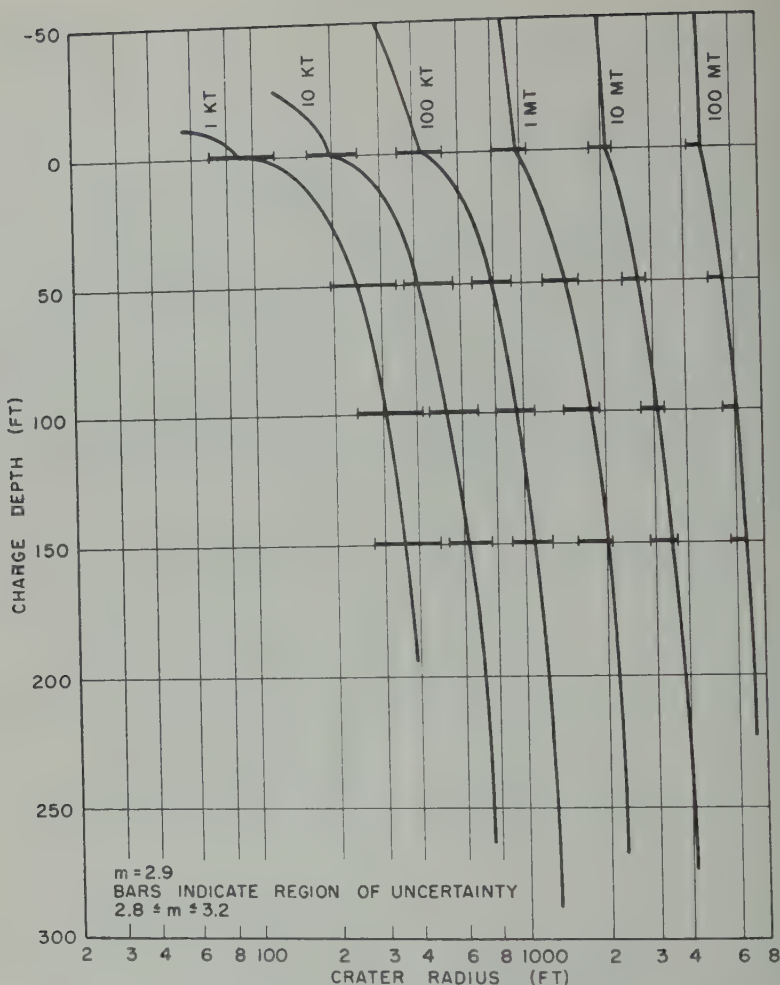


Fig. 14. Crater radius vs. charge depth, dry clay (Fig. 4.7, WT-920).

and is used only for assistance in relatively minor aspects. In adopting an empirical approach, it would, of course, be absurd to ignore the information, however meager, in regard to the physical mechanism and particularly in the distinction between the mechanisms occurring in TNT and in nuclear explosions. On the other hand, too much dependence on cube-root scaling is likely to give the illusion of a precision in prediction unjustified by the facts.

The development described below was undertaken within the framework that the desirable result from a military standpoint is the construction of graphical or analytical relations such that knowledge of the yield, soil, and depth will permit each prediction of the crater dimensions.

It is postulated that the shape of a crater for the craters of interest is primarily dependent on its size, and hence the first attempt is to predict crater radius in terms of the three parameters just mentioned, with the expectation that later analysis can be made to predict depth and other shape aspects once the radius prediction has been accomplished.

Development of the extrapolation method. It was decided to study first the effect of soil type, second the effect of depth, and third the effect of yield. In looking at the available information it was at once apparent that in regard to both soil type and depth the data on megaton explosions are useless, since all these shots were fired at one depth (essentially zero) and in

the ('coral' atoll); hence, it was finally decided that the germane approach appeared to look first only at TNT data and from that to establish an extrapolation procedure, to adjust the values of the parameters that the Jangle U and Jangle S shots were consistent; and, finally, to investigate the validity of the procedure and compare the results with the measurements of nuclear craters at the Marshalls.

For a soil is an appropriate one to look at since there are considerable amounts of explosive data and data from two nuclear craters for that soil, data are available in the range $\lambda_c = -0.13$ to $+1.0$. Within this range greatest variation lies in the neighborhood of $\lambda_c = 0.14$. Data on the TNT shots of this scaled depth plotted in Figure 8, which shows crater radius plotted against yield on log paper both in Figure 9 is a similar plot for data on TNT craters of depth $\lambda_c = 0.50$ and $\lambda_c = -0.14$

(minus indicates above the surface). The scatter of the points shown on these graphs is typical of the scatter shown in every case where several essentially identical shots have been fired. It is conservative to say that the uncertainty in the value of radius for any specific combination of soil type, charge size, and charge depth is at least 10 per cent. Consequently, the plus and minus 10 per cent limits at the maximum and minimum charge sizes shown here are marked on Figure 8. For extrapolation purposes, the reciprocal slope, m , of the most probable line is 3.4. (The actual value measured on the graph is 3.39. However, since the second figure is probably of somewhat doubtful validity, all such numbers are rounded off to two figures.) To permit an estimate of the uncertainty in extrapolation, maximum and minimum slopes within the 10 per cent uncertainty just mentioned have also been plotted. These slopes are $m = 3.0$ and $m = 4.1$. This elementary analysis has been

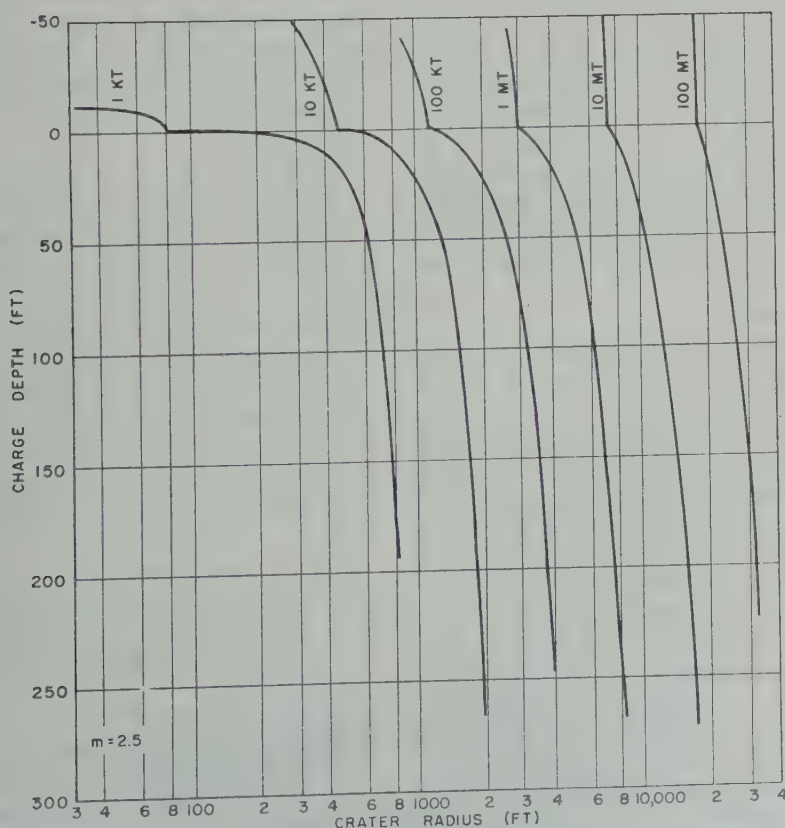


Fig. 15. Crater radius vs. charge depth, wet clay (Fig. 4.8, WT-920).

undertaken with the data on Figure 8 only, and lines of the slopes so determined have then been drawn on Figure 9. The analysis has been limited to Figure 8, both because the scaled depth $\lambda_c = 0.14$ is of major interest and also because a greater range of yields for TNT shots is available for this scaled depth than for any other.

It is apparent that m , the reciprocal of the slope when crater radius is plotted against yield on a log-log basis, is related to R and W as

$$R = KW^{1/m}$$

In the remainder of the report m is referred to as the 'scaling exponent.'

Using the best-fit value for m , 3.4, and the experimental data of Tables A4 and A6, the solid line of Figure 10 has been constructed. On this figure the scaled radius (on the basis $m = 3.4$) is plotted against the scaled charge depth (on the basis $m = 3$). (Since the range of scaled depths is small in the interval of greatest interest, the distinction between determining scaled depths on the basis $m = 3.0$ and on the basis $m = 3.4$ is relatively trivial and will not affect the conclusions reached in this analysis.)

The next step is the determination of the curve for nuclear charges based on this curve for TNT charges. In this procedure consideration must be given to the difference in mechanism of nuclear and TNT bursts, particularly for bursts on the surface or at very low heights above the surface.

In the early stages of a nuclear explosion fired at or near the interface between air and earth, the shock-wave velocity is very much higher in the air than in the earth. (Griggs [1952], in predicting the effects of Jangle U, computes shock-wave velocities in air to be approximately 25 times those in soil in the radius range from approximately $\lambda = 0.1$ to $\lambda = 1.0$. Similarly, Porzel [1952], in predicting the effects of Ivy Mike, estimates shock velocities in the air and water-soaked sand for high overpressures such that in the early stages of a nuclear explosion the ratio of velocity in air to velocity in soil may be as high as 1000 : 1.) At a time when the nuclear explosion process has proceeded to the point where the average energy density (the total energy contained within the shock wave divided by the total volume within it) within

the boundary of the shock wave is equal to the average energy density at the surface of a spherical TNT charge which has been detonated at its center, the envelope of the nuclear explosion is essentially hemispherical. If average energy density is a good criterion of crater size and shape, the crater formed by a given nuclear energy release on the surface should be similar to the crater formed by a TNT charge of the same yield fired well above the surface. (Actually, as Porzel points out, at a time when the nuclear shock wave has reached the same radius as that of the TNT sphere of equivalent energy released, and hence when average energy densities are equal, there is still an enormous difference in the two situations, since the mass enclosed within the shock wave in the TNT explosion is some 1500 times that in the nuclear explosion. Hence, in the nuclear situation the pressures are very much higher and the durations shorter than in the TNT situation.) The crater resulting from a nuclear surface charge should differ extensively from that produced by a TNT charge whose center of gravity is at the surface, both because of the different mechanism mentioned above and because a hemispherical excavation was required before the TNT charge could be placed.

Consider a nuclear charge at $\lambda_c = -0.13$. Within its shock wave the total energy will be identically the same as that within a sphere of TNT tangent to the surface when both shock waves reach the surface. This argument can be summarized by saying that the crater radius produced by a low aboveground nuclear shot should be essentially independent of height, and (if the efficiency were 100 per cent) should have about the same value as that produced by a TNT shot at $\lambda_c = -0.13$. On this basis the dotted curve in the region AB has been drawn in Figure 10.

Since the energy partition in the two types of explosion is significantly different, particularly in the roughly 15 per cent of the yield of the nuclear explosion which takes the form of prompt radiation, it seems necessary to consider an equivalence factor of less than 1 for the cratering effects of nuclear explosions. The experimental evidence on this point is very meager, being limited to the Jangle S and Jangle U shots. The data from these two shots can be placed on this graph with the equivalence

as a parameter. Thus segment *DE* on 11 represents Jangle S shot for a radio- yield of 1.2 kt times the factors shown line, with radius scaled on the basis of 3. y the segment *FG* represents Jangle U the basis 1.2 kt times the factors shown following the same procedure. The inter- of these segments with the solid curve plies equivalence factors of 0.14 and 1.0 gle S and Jangle U, respectively. It is ant not to attach too much precision to uments, for the following reasons: (1) the ata have an important scatter amounting

to an uncertainty of the order of ± 10 per cent in scaled radius; (2) the TNT curve in the neighborhood of zero charge depth has an additional uncertainty, inasmuch as the radius of the TNT sphere is 0.13 on the basis $m = 3$.

Consideration of the difference in energy partition leads then to the qualitative conclusion that there should be gross differences in sensitivity of crater radius to changes in the charge height in the immediate vicinity of the surface between nuclear and chemical explosives. These differences should in general be that a nuclear charge placed slightly above surface will

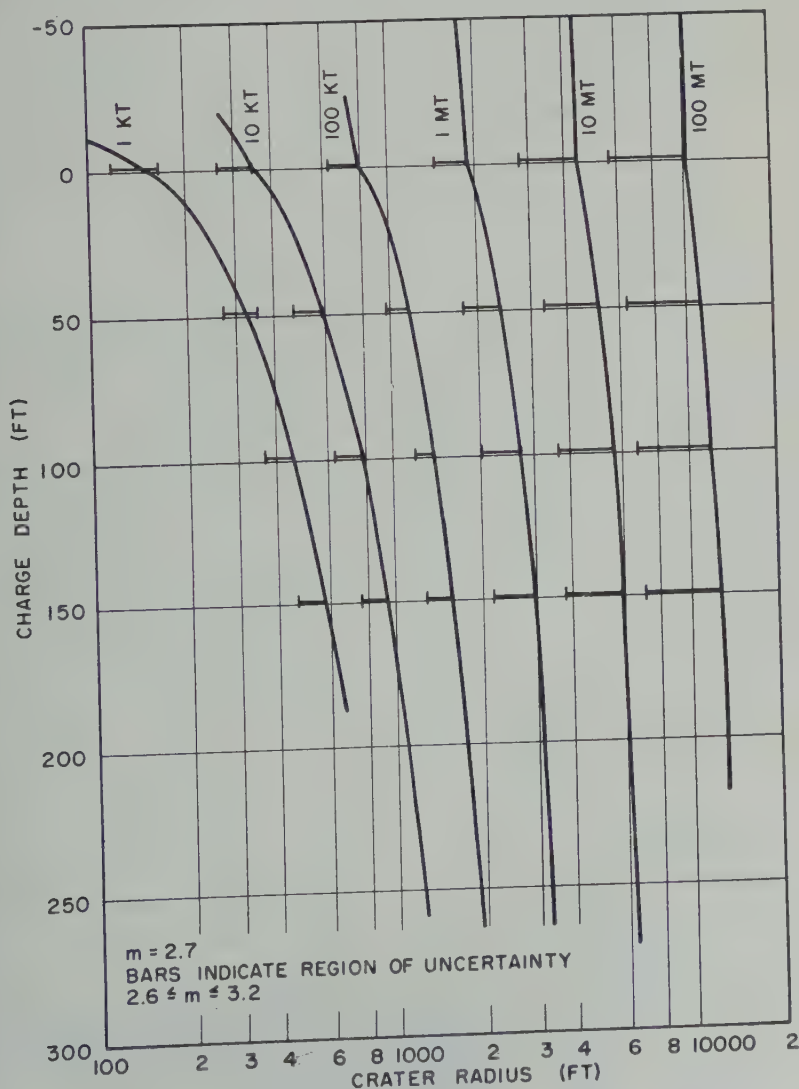


Fig. 16. Crater radius vs. charge depth, dry sand (Fig. 4.9, WT-920).

produce a crater relatively insensitive to further changes in height, whereas a nuclear charge slightly below the surface will have a crater radius extremely sensitive to further changes in depth. Thus the actual scaled crater radius to be expected from a nuclear explosion probably falls on the dashed curve *ABC*. This has been drawn through the point representing 1.0 effectiveness factor for Jangle S and may well be a more fruitful method of thinking of crater predictions from nuclear explosions than attention to equivalence factor and its variation with height or depth.

The procedure described for constructing both the TNT and the nuclear curves shown in Figure 10 can be performed equally well using values of *m* other than the best-fit value 3.4. Other appropriate values of *m* as indicated in Figure 8 are 3.0, representing both conventional cube-root scaling and the lower limit of slope on the basis of the 10 per cent uncertainty in experimental values postulated earlier, and 4.1 representing the upper limit. The two curves have been plotted together in Figure 11. Note that when the dashed curve *ABC* for these values of *m* is drawn through the point repre-

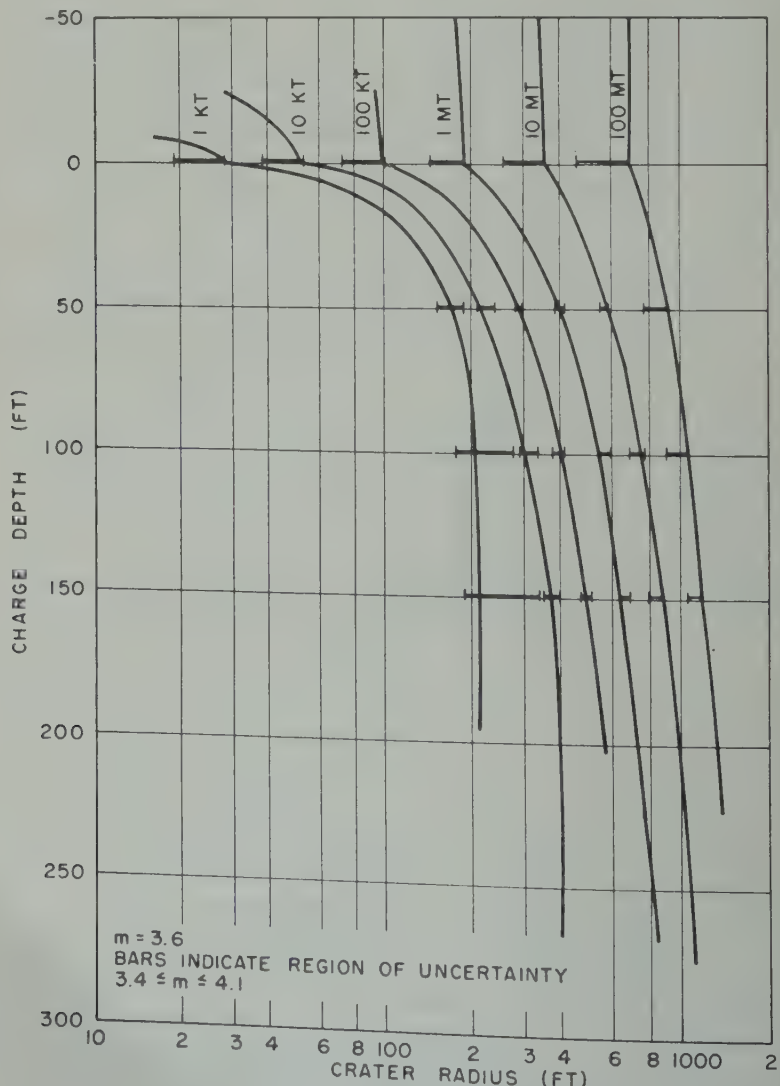
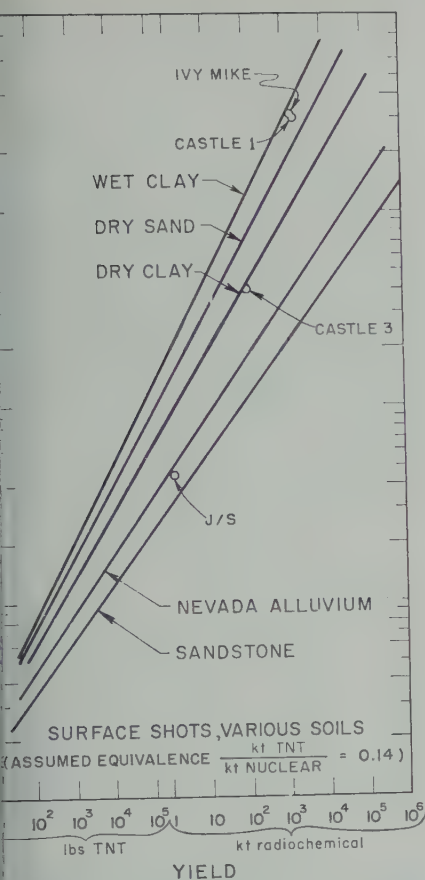


Fig. 17. Crater radius vs. charge depth, sandstone (Fig. 4.10, WT-920).

TABLE 1

<i>m</i>	Equivalence Factors		
	Jangle S	Surface	Jangle U
3.0	0.07	0.03	0.35
3.4	0.14	0.08	1.0
4.1	0.50	0.31	10.0



extrapolation should be available in the simplest possible form for quick use without computation, the nuclear curves shown in Figures 10 and 11 have been replotted in the form of radius in feet against charge depth in feet, with yield as a parameter. This has been done in Figure 12, in which for each yield shown both the most probable value ($m = 3.4$) and the limiting values $m = 3.0$ and 4.1 are shown.

The estimates for this soil for the most probable value of the scaling exponent m (3.4) are replotted in Figure 13. Range of uncertainty ($m = 3.0$ and $m = 4.1$) is indicated by short horizontal bars attached to each of the parametric yield curves.

The same kind of analysis has been carried through for dry clay, dry sand, wet clay, and sandstone; the results of these analyses are included in Figures 14 through 17. For these other soils no nuclear data are available, and hence the efficiencies found in the Nevada soil have been used in the following fashion. For the most probable value of the scaling exponent m in each of these other soils, the variation of equivalence with depth at Nevada for $m = 3.4$ has been used. Similarly, for the lowest value of m for each of these other soils the same variation of efficiency with depth has been used as was found at Nevada for the lowest value of m there, namely, 3.0. The corresponding analysis

0 equivalence factor for Jangle S the re a little unrealistic in that the curve $m = 3$ implies a larger crater for an ound nuclear shot than seems reasonable, e AB for $m = 4.1$ shows a smaller value ms reasonable.

equivalence factors for Jangle S and J as shown in Figures 10 and 11 are ensitive functions of parameter m . It is be noted that if Jangle S had been t at the surface the efficiency factor ave been significantly different because ute angle of the TNT curve as it crosses depth axis. Table 1 indicates the values equivalence factor as a function of m for both vents and also the estimated value for a e Jangle S except one fired precisely arface.

for military purposes, the data for

TABLE 2. Scaling Exponent, m , for Several Soils

Soil	Most Probable	Minimum	Maximum
Nevada	3.4	3.0	4.1
Dry clay	2.9	2.8	3.2
Wet clay	2.5	2.0	3.3
Dry sand	2.7	2.6	3.2
Sandstone	3.6	3.4	4.1

has been made for the upper limiting value of m . The most probable and limiting values of m for all the soils reported here are listed in Table 2. In each, the available data have been plotted in the same form as was shown in Figures 8 and 9, the best straight line was drawn for those points, and then values of radius 10 per cent above and below the curve were marked at the upper and lower limits of the charge sizes considered. (TNT data from charges less than 200 lb were not reviewed.) By this procedure, the limiting values of m have the greatest range for soils in which no large TNT charges have been fired; this is appropriate, since in fact the extrapolation is less certain in such cases.

For wet clay, Figure 15, so few TNT data are available that crater radius has been predicted only for the most probable value of the scaling exponent m .

In Figure 18 the results for surface charges in various soils are shown. For each soil the line drawn is that for the most probable value of m . On this curve also are shown the nuclear craters for surface shots at Nevada and in the Marshalls. In plotting the results of the nuclear explosions in this figure, the equivalence factor found for a shot precisely at the surface for the scaling exponent $m = 3.4$ has been assumed to be applicable to the explosions in the Pacific. The logarithmic grid has been adjusted in the region of 1 kt to include this equivalence for all larger yields. Hence the graph can be entered directly with the value of radiochemical yield. This graph gives a realistic indication of the uncertainty in crater prediction depending on the properties of the soil.

All data that have been used in the development of the extrapolation method presented here are summarized in the Appendix. The Appendix also includes data for some TNT shots, namely, those in wet sand, as well as some nuclear charges, such as Trinity, which were not used in the actual analyses presented here. The wet sand TNT results were not used because data on only one charge size was found and hence a value of slope could not be established. A value for Trinity was not used because the scaled height is greater than that of interest in this report.

Comments on the extrapolation method. It should be noted explicitly that the extrapolation method described here is based on an empirical

equation of the form

$$R = f(W, m) \cdot f(\lambda_c)$$

or

$$R = (WE)^{1/m} \cdot f(\lambda_c)$$

where E is an efficiency which depends on medium, scaled charge depth, and type of explosive. This is not the only form of equation that can be postulated, and defended. The available data are so meager, and their scatter around the curve representing any specific equation is so great, that it is not possible to present to establish unequivocally the relative validity of alternative forms of the empirical equation.

The suggestion has been made that an equation of the form

$$R = (WE)^{1/3} \cdot f(\lambda_c) \cdot f(m)$$

is more satisfactory.

One piece of information that has been put forward as favoring equation 6 is the result of some cratering experiments in the Marshall Islands. These experiments were run under the direction of Dr. H. Kirk Stephenson, currently on the staff of the National Science Foundation. Quoting from Memorandum SWPEF 2/9.2 (354.2) dated Nov. 26, 1954:

1. A series of high explosive shots were fired at Eniwetok (Flora) Island, Eniwetok Atoll, in the spring of 1952. These shots consisted of a combination of R-7HDA(c-2)R-7-HCA (Tetrytol), priming cord, and blasting caps piled in a beehive shape on the surface which had been excavated down to the high-tide level. A dike was established around the charge to prevent wave interference but this proved ineffective. In addition to seismic shock information, the crater radii were determined. The crater data obtained from these HE shots at the Pacific Proving Grounds may be used to establish a soil factor for comparing saturated coral with Nevada soil. The data are summarized in the table.

W, tons TNT equivalent	Scale	Crater Radius R_c , ft	$R_c/W^{1/3}$
	Height to c.g., λ_c		
1	12.6	0.06	27.5
5	21.5	0.06	32
10	27.1	0.06	37.5
15	31.1	0.06	45.5
20	34.2	0.06	50
Over-all average			1.60
Average if first shot omitted			1.46

responding high-explosive data from taken from Tables A4 and A6 give a $R_c/W^{1/3}$ of about 0.08. If equation 6 is it is assumed that the effect of soil is nt of the effect of charge size, craters arshalls should be expected to be 1.8 es as large (in radius) as craters from charge sizes and depths in Nevada. nilar manner it is found that the value $1/3$ for megaton surface shots in the is about 1.0, while that for the kiloton ot in Nevada is 0.34, which implies rshall craters will be some 3 times an Nevada craters. Actually, if the finite value of $D_c/W^{1/3}$ is taken into particularly for Jangle S shot, the suggests that scaled crater radii for charges in the Marshalls are twice as for those in Nevada. Since this is the re that was obtained for high-explosive t is tempting and not implausible to all scaled crater radii in the Marshalls ery close to twice those in Nevada. gh the precise data quoted from the memorandum were not at hand during opment of the extrapolation method in the previous section, some prior n of them was held with Dr. Stephenson hone. At that time it was Dr. Stephen- lief that the data themselves were some- reliable because all the craters were ashed before measurement. In addition s improper to assume that the charac- for cratering purposes, of the water- d coral sand involved in the high- e tests are identical with the charac- of the more coherent water-saturated ck involved in the nuclear shots. her equation 4, 5, or 6 is used for the ation of TNT data to megaton nuclear ns, a certain range of uncertainty in R yn, resulting from reasonable values il for the uncertainty in $f(\lambda_c)$ and $f(m)$. ertainty in R shown by equation 6 is than that shown by equation 3. her and more important benefit adduced

for equation 6 is that the predicted crater radii for megaton explosions have a smaller spread when soil characteristics are changed.

It is the opinion of the author that the benefits indicated are illusory and that equation 3 has a slightly better basis. The true value of crater radius produced by a megaton explosion in any medium other than that existing in the Marshall Islands will remain unknown until such a shot is fired and the resulting crater measured. In the meantime, caution in stating the expected values and their uncertainties is of vastly greater military use than overoptimism.

REFERENCES

- Bishop, J. A., and F. E. Lowance, Physical characteristics of crater and lip, Naval Civ. Eng. Research & Eval. Lab., *Operation Jangle Rept. WT-399* (Secret), 1952.
- Campbell, D. C., Some HE tests and observations on craters and base surges, AFSWP, *Operation Jangle Rept. WT-410* (Secret), Nov. 1, 1951.
- Gaylord, J. L., Photographic crater survey, Lookout Mtn. Lab., *Operation Ivy Rept. WT-618 (SRD)*, 1952.
- Griggs, D. T., Notes on surface and underground bursts, AFSWP, *Operation Jangle Rept. WT-378* (Secret), 1952.
- Mem. SWPEF 2/924 (354.2)*, 1954.
- Perret, W. R., Cratering produced by nuclear weapons, *Sandia Corp. Tech. Mem. 1922-2(23)*, Jan. 2, 1954.
- Porzel, F. B., Soil pressures and energy transfer on Mike shot, *Los Alamos Sci. Lab. Rept. LA-1529* (Secret), 1952.
- Swift, L. M., and D. C. Sachs, *Small Explosion Tests—Phase II of Project Mole*, Stanford Research Institute, May 1954.
- Underground Explosion Test Program, *Final Rept.*, vol. I, Engineering Research Associates, Aug. 30, 1952.
- Underground Explosion Test Program, *Tech. Rept. 4*, Granite and Limestone, *Tech. Rept. 5*, Sandstone, vol. I, Engineering Research Associates, Feb. 15, 1953.
- Vaile, R. B., Jr., Crater survey, *Operation Castle Rept. WT-920*, 1955.
- Vaile, R. B., Jr., *Small Explosion Tests—Phase I of Project Mole*, Stanford Research Institute, January 1953.

(Manuscript received July 10, 1961.)

APPENDIX
SUMMARY OF AVAILABLE CRATER DATA

TABLE A1. Nuclear Crater Measurements*

Shot	Soil	RC Yield	Height of Burst		Crater Radius†		Crater Depth‡	
			ft	λ	ft	λ	ft	λ
Trinity	Dry sand	~20 kt	100		550		9.5	
Greenhouse	Saturated coral							
Easy‡	sand	46.7 kt	300	-0.664	418	0.925	2.4	0.008
Jangle S	Desert alluvium	1.2 kt	3.5		45	0.336	17	0.127
Jangle U	Desert alluvium	1.2 kt	-17	0.127	129	0.961	53	0.396
Ivy Mike‡	Saturated coral							
	sand	~14 Mt	35		3120		164	
					(2800)§			
Castle 1	Saturated coral							
	sand	~15 Mt	7		3000		240	
Castle 3	Saturated coral							
	sand	~100 kt	13.6		400		75	
Teapot Ess	Desert alluvium		-70		147		90	

* All data except Castle and Teapot data are obtained from *Perret* [1954].
† All crater radii are measured at original ground level.
‡ Owing to scour from water rushing back in, and to aging (for Greenhouse), measured diameters may be large by 10 to 30 per cent, and measured apparent crater depths may be shallow by a factor of 2 or more.
§ In Memorandum SWPEF 2/924 (354.2) dated Nov. 26, 1954, the statement is made that plotting the Ivy Mike data on an expanded vertical scale gives a value for crater radius of 2800 ft ($\lambda = 1.02$).

TABLE A2. TNT Crater Measurements in Dry Sand, Dry Clay, and Wet Clay*
Underground Explosion Test Program
Site: Dugway Proving Grounds

Round	Charge Weight, lb TNT	Charge Depth		Crater Radius†		Crater Depth	
		ft	λ	ft	λ	ft	λ
101	320	-3.5	-0.51	4	0.59	0.5	0.07
102	320	0.0	0.0	7.68	1.12	2.5	0.37
103	320	1.3	0.19	10.88	1.59	6	0.88
104	320	3.5	0.51	12	1.75	6.5	0.95
105	320	7.0	1.02	15.5	2.26	8.5	1.24
106	320	14.0	2.04	16.75	2.45	4.5	0.66
107	320	21.0	3.07	13.5	1.97	3.5	0.51
108	2,560	2.6	0.19	19	1.39	9.75	0.71
109	2,560	7.0	0.51	24.75	1.81	8.5	0.62
110	320	3.5	0.51	13	1.9	7.5	1.10
111	8	2.5	1.25	6	3	4	2
112	2,560	7.0	0.51	30	2.2	12	0.88
113	320	3.5	0.51	14	2.0	6.75	0.99
114	8	2.5	1.25	6	3	3.5	1.75
115	40,000	17.5	0.51	75	2.19	23	0.67
116	320	8.75	1.28	18.5	2.7	9	1.32
301	320	-3.5	-0.51	2.5	0.37	1	0.15
302	320	0.0	0.00	7.25	1.06	4	0.58
303	320	1.3	0.19	9	1.3	5.5	0.80
304	320	3.5	0.51	10.5	1.5	6	0.88
305	320	7.0	1.02	11.75	1.72	7	1.02
306	320	14.0	2.04	15	2.2	1	0.15
307	320	21.0	3.07	10	1.46	1	0.15
308	2,560	2.6	0.19	20	1.46	12	0.88
309	2,560	7.0	0.51	21.5	1.57	15.5	1.13
310	320	3.5	0.51	11	1.6	7	1.02
311	8	2.0	1.0	4	2	2.5	1.25
312	2,560	7.0	0.51	26	1.90	15	1.09
313	320	3.5	0.51	12.75	1.86	8	1.17
314	8	2.5	1.25	4.5	2.25	3	1.5
315	40,000	17.5	0.51	64	1.87	42	1.23
316	110	2.45	0.51	9	1.87	6	1.25
317	2,560	7.0	0.51	23	1.68	15.5	1.13
318	320,000	35.0	0.51	120	1.75	60	0.88
319	2,560	7.0	0.51	23	1.68	13.5	0.98
Sym.	320	7.0	1.02	12.5	1.83	7	1.02
401	8	2.5	1.25	7	3.5	5	2.5
402	320	2.5	0.36	18.75	2.74	10	1.46
403	2,560	5.0	0.36	41.75	3.05	12.75	0.93
404	320	2.5	0.36	17.5	2.56	11.5	1.68
405	8	2.5	1.25	6	3	4.1	2.05

*Taken from Appendix G, *Underground Explosion Test Program* [1952]
†Crater radii are measured at original ground level.

TABLE A3. TNT Crater Measurements in Limestone, Granite, and Sandstone*
Underground Explosion Test Program
Site: Dugway Proving Ground

Soil	Round	Charge Weight, lb TNT	Charge Depth		Crater Radius†		Crater Depth‡	
			ft	λ	ft	λ	ft	λ
Limestone and granite	501	320	6.6	0.97	11.2	1.64	9.1	1.33
	502	320	2.5	0.365	8.3	1.21	3.9	0.51
	601	320	-2.5	-0.365	1.20	0.175
	602	320	0.0	0.00	8.43	1.23	1.7	0.22
	603	320	2.5	0.365	9.70	1.42	2.6	0.33
	604	320	5.0	0.73	14.5	2.12	5.0	0.73
	605	320	12.5	1.83	17.1	2.50	6.1	0.80
	606	320	25.0	3.65	5.20	0.76	2.0	0.26
	607	320	2.5	0.365	14.4	2.11	5.3	0.70
	608	320	2.5	0.365	14.0	2.05	4.6	0.62
	609	2,560	5.0	0.365	25.2	1.84	10.2	0.79
	610	2,560	5.0	0.365	23.1	1.69	8.7	0.63
Sandstone	611	320	2.5	0.365	13.4	1.96	5.0	0.73
	612	320	17.0	2.49	13.2	1.93	7.6	1.11
	801	320	-2.5	-0.365	0.0	0.0	0.0	0.00
	802	320	0.0	0.0	5.6	0.82	2.3	0.33
	803	320	2.5	0.365	11.6	1.69	4.8	0.70
	804	320	5.0	0.73	14.0	2.04	7.6	1.11
	805	320	12.5	1.82	9.3	1.36	14.9	2.15
	806	320	25.0	3.65	0.0	0.00	§	§
	807	320	2.5	0.365	14.3	2.09	5.1	0.73
	808	320	2.5	0.365	13.1	1.91	5.8	0.83
	809	1,080	3.75	0.365	19.0	1.85	8.6	0.84
	810	2,560	5.0	0.365	32.6	2.38	9.7	0.71
	811	2,560	5.0	0.365	25.1	1.83	10.5	0.77
	812	2,560	5.0	0.365	23.3	1.70	11.0	0.80
	813	10,000	7.9	0.365	39.4	1.83	16.1	0.73
	814	40,000	12.5	0.365	56.5	1.65	26.9	0.73
	815	40,000	12.5	0.365	70.5	2.06	26.9	0.79
	816	40,000	12.5	0.365	53.6	1.56¶	27.5	0.80
	817	320,000	25.0	0.365	94.8	1.38¶	47.0	0.69
	818	320	2.5	0.365	17.5	2.56	6.0	0.88
	819	320	2.5	0.365	15.6	2.28	6.5	0.93

* Obtained from *Underground Explosion Test Program* [1953].
† All crater radii are measured at original ground level.
‡ Average crater depth (D_k) is the average of the measurements of the vertical distance from the deepest point of the crater, not necessarily directly under the charge, to the surface, one measurement being made on each of the four vertical sections available for each crater. This depth is not significant unless the deepest point is below the bottom of the excavation made to place the charge. The charge hole was obliterated by all the detonations at the sandstone site except round 306.
§ The damage did not extend to the surface and is not comparable with other rounds; the sides of the original charge hole were damaged up to an average slant distance of 5.6 ft from the center of gravity of the charge.
^{||} Crater shape was estimated; the breakthrough volume is not included.
¶ Average of eight measurements scaled from the vertical crater sections.

TABLE A4. TNT Crater Measurements in Desert Alluvium*
 Operation: Jangle HE Shots
 Site: Nevada Proving Grounds (Yucca Flat)

Charge Weight, lb of TNT	Charge Depth		Crater Radius†		Crater Depth	
	ft	λ	ft	λ	ft	λ
2,560	2.01	0.15	18.2	1.33	6.5	0.47
40,000	4.63	0.15	38.6	1.13	14.9	0.44
2,560	6.79	0.50	19.8	1.45	10.8	0.79
2,560	-2.01	-0.15	6.4	0.47	1.9	0.14
2,560	4.02	0.30	19.6	1.43	7.8	0.57
2,560	3.00	0.22	19.7	1.44	6.7	0.49
2,560	2.58	0.19	18.9	1.38	6.9	0.50
216	1.08	0.18	†	†	†	†
216	0.83	0.14	8.2	1.37	3.5	0.58
216	3.00	0.50	11.3	1.88	5.5	0.92

ined from *Campbell* [1951].

crater radii are measured at original ground level.

ial detonation.

alts from a corresponding 177-lb Pentolite charge are not included in this summary.

TABLE A5. TNT Crater Measurements in Dry Clay*
 Project: Mole (Stanford Research Institute)
 Site: Dugway Proving Grounds

Charge Weight, lb of TNT	Charge Depth		Crater Radius†		Crater Depth	
	ft	λ	ft	λ	ft	λ
256	6.35	1.00	11.1	1.73	5.5	0.86
256	6.35	1.00	10.9	1.72	6.0	0.94
256	3.18	0.50	10.5	1.65	6.3	0.99
256	3.18	0.50	9.5	1.50	5.4	0.85
256	1.65	0.26	9.1	1.43	6.2	0.98
256	0.0	0.00	6.6	1.04	3.9	0.61
256	-0.83	-0.13	4.4	0.69	1.5	0.24

ained from *Vaile* [1953].

crater radii are measured at original ground level.

TABLE A6. TNT Crater Measurements in Desert Alluvium*
 Project: Mole (Stanford Research Institute)
 Site: Nevada Proving Grounds (Yucca Flat)

Round	Charge Weight, lb of TNT	Charge Depth		Crater Radius†		Crater Depth	
		ft	λ	ft	λ	ft	λ
202	256	6.35	1.00	11.5	1.81	5.7	0.9
212	256	6.35	1.00	10.7	1.69	6.1	0.9
203	256	3.18	0.50	8.4	1.32	4.0	0.6
204	256	1.65	0.26	9.2	1.45	2.9	0.4
205	256	0.83	0.13	8.8	1.39	2.5	0.3
206	256	0.0	0.00	6.4	1.01	1.9	0.3
207	256	-0.83	-0.13	3.5	0.55	1.4	0.2

* Obtained from *Vaile* [1953].

† All crater radii are measured at original ground level.

TABLE A7. TNT Crater Measurements in Wet Sand*
 Project: Mole (Stanford Research Institute)
 Site: Camp Cooke, California

Round	Charge Weight, lb of TNT	Charge Depth		Crater Radius†		Crater Depth	
		ft	λ	ft	λ	ft	λ
304‡	256	4.83	0.75	18.6	2.94‡	6.6‡	1.04‡
301	256	3.18	0.50	19.1	3.01
302	256	3.18	0.50	19.9	3.14	6.3	0.99
309	256	3.18	0.50	15.6	2.45	6.1	0.96
310	256	3.18	0.50	16.8	2.64	5.2	0.82
305	256	1.65	0.26	14.3	2.26	6.3	0.99
306	256	0.83	0.13	12.8	2.01	3.7	0.58
307	256	0.00	0.00	10.2	1.61	4.8	0.75
308	256	-0.83	-0.13	8.8	1.39	4.0	0.63

* Obtained from *Swift and Sachs* [1954].

† All crater radii are measured at original ground level.

‡ Round 304 was shot in the crater of round 303.

TABLE A8. TNT Crater Measurements in Wet Clay*
 Project: Mole (Stanford Research Institute)
 Site: Camp Cooke, California

Round	Charge Weight, lb of TNT	Charge Depth		Crater Radius†		Crater Depth	
		ft	λ	ft	λ	ft	λ
311	256	3.18	0.50	15.5	2.45	11.2	1.70
312	256	3.18	0.50	17.8	2.80	9.0	1.42
313	256	-0.83	-0.13	5.8	0.91	3.4	0.53

* Obtained from *Swift and Sachs* [1954].

† All crater radii are measured at original ground level.

Nuclear Craters and Preliminary Theory of the Mechanics of Explosive Crater Formation¹

M. D. NORDYKE

*Lawrence Radiation Laboratory, University of California
Livermore, California*

Abstract. Four nuclear craters have been produced at the Nevada test site. Three were from kiloton nuclear explosions in desert alluvium, a sand-gravel mix, and the fourth was from a 15-ton nuclear explosion beneath the sloping side of a bedded tuff mesa. Comparison of these craters with high-explosive craters in alluvium shows that, within experimental error, the craters produced by subsurface nuclear explosions are quite comparable with those produced by equal chemical explosions. Experimental data from these nuclear and chemical explosive cratering programs and theoretical machine calculations of the behavior of underground explosives make it possible to construct a picture of the major mechanisms that contribute to the formation of explosive craters. These mechanisms include compaction and plastic deformation of the medium immediately surrounding the explosion, spalling of the surface above the explosion by the tensile wave generated at the free surface of the ground, and acceleration of the fractured material overlying the explosion cavity by the gases trapped in the cavity, before and during their escape. The role that each mechanism plays changes with the scaled depth of burst of the explosive and to some extent with the material. The contribution that each makes is outlined for four typical craters representing shallow, optimum, and deep burial depths. For surface burial, plastic deformation and compaction are the principal actions; for shallow burial depth, spall is the dominant feature; for optimum depth, gas acceleration becomes the most important mechanism; and for deep burial, subsidence into the cavity produced by plastic deformation and compaction is the major factor. Differences to be expected between explosion craters and craters resulting from impact explosions such as those produced by meteors are examined. The relative contribution of each of these mechanisms is also estimated for apparent crater depth vs. depth of burst.

In recent years a large number of cratering experiments have been performed, both nuclear and chemical, accompanied by field experiments and to answer many questions about the mechanisms of cratering. Many of the interesting chemical cratering programs appear in the following paper [Murphey and Vortman, 1961]. In this paper I shall briefly outline the data obtained from the nuclear craters at the Nevada test site and present a preliminary theory of the mechanisms of explosive cratering based on the results of the various field programs.

NUCLEAR CRATERS

Four nuclear explosives have been detonated at the Nevada test site (NTS) under conditions that resulted in the creation of large craters. Three of these were fired for the purposes of nuclear

weapons effects studies and hence were at rather shallow depths of burst, much shallower than proposed Plowshare applications but right in the region of interest for meteoritic impact cratering explosions. All three of these detonations occurred in the valley alluvial fill of area 10 at NTS, a medium characterized as a loose, sand-gravel mix with a density of approximately 1.5 to 1.7 and a water content at depth of about 10 per cent. The fourth nuclear crater, Neptune, was made in the bedded tuff of the Rainier mesa at NTS. This medium is a weakly cemented volcanic ash in which all the deep underground nuclear explosions (Rainier, Blanca, Logan, and others) have been fired [Johnson, Higgins, and Violett, 1959].

In Table 1 are shown the apparent crater dimensions, measured with respect to the original ground level for these four nuclear craters. In Figure 1 the crater profiles have all been scaled to 1 kiloton for comparison purposes. Some discussion of each crater will be useful.

Jangle S. In the Jangle S event a 1.2-kt

This paper is a shortened version of two papers presented at the Geophysical Laboratory-Livermore Radiation Laboratory Cratering Symposium held at the Geophysical Laboratory, in Washington, D. C., on March 28 and 29, 1961.

TABLE 1. Summary of Data from Nuclear Craters at the Nevada Test Site

Shot Name	Jangle S	Jangle U	Teapot Ess	Neptune
Medium	Alluvium	Alluvium	Alluvium	Tuff
Yield, kt	1.2	1.2	1.2	0.115
Depth of burst, ft	-3.5*	17	67	100†
Apparent crater radius R, ft	45	130	146	100
Apparent crater depth, D, ft	21	53	90	35
Apparent crater volume, yd ³	1650	3.7×10^4	9.6×10^4	2.2×10^4
R/D	2.15	2.45	1.62	2.86
Lip height, ft	...	8	19	...

* Detonated 3.5 ft above surface.
† Neptune was detonated 100 ft beneath a 30° slope.

nuclear explosive was detonated 3.5 feet above the surface of the ground in late 1951. As can be seen from Figure 1, the crater formed was very small. Essentially no loose fallback material was found in the crater. The crater and lip were formed almost entirely by plastic deformation of the ground by the action of the fireball. An additional reason for the small size is that, for a nuclear device, about a third of the released energy is in the form of thermal and X-ray radiation, which is lost immediately for a surface burst. For a subsurface burst, this radiation energy vaporizes and melts the medium surrounding the device, and so part of it is available for later utilization.

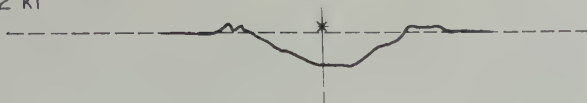
Jangle U. The 1.2-kt nuclear device used for the Jangle U event, also detonated in late 1951, was placed in a concrete-lined room, 10 by 10 by 8 feet high. The hole from the top of the room to the surface was stemmed with a sandbag plug. The center of the device was 17 feet below the surface of the ground. The crater resulting from this explosion was considerably larger than the Jangle S crater but still much smaller than the maximum possible crater for a 1.2-kt explosion. A brief flash of the fireball was observed, but for a much shorter time than at Jangle S. A dense dust cloud rose to a height of about 6000 feet, and a base surge spread out radially to a distance of about 1 mile. Almost all the radioactivity escaped to the atmosphere and was deposited on the surface within 10 miles.

Teapot Ess. Teapot Ess was fired during Operation Teapot in March 1955, at a site very near the Jangle U crater. The 1.2-kt device was

located 67 feet below the surface, at the bottom of a 10-foot-diameter hole. A 30-foot-diameter hole was provided for personnel access before the explosion. The device was packed closely with sandbag plugs and both holes were filled with loose alluvial material before the detonation. A flash of very short duration was also observed on detonation of this device. Again a base surge about 1 mile in radius was formed. Survey of the total radioactivity after the explosion revealed that about 70 per cent of it was released in a manner similar to that of Jangle U. The crater produced was considerably larger than the Jangle U crater because of the greater depth of burst, which gave much better coupling of the explosion energy to the ground. On the basis of high-explosive experimental data at large depths of burst, however, it was still much smaller than the maximum possible from a 1.2-kt device.

An extensive program to delineate the true crater was undertaken for the Teapot event. Twenty-one colored sand columns were erected, as shown in Figure 2, along a diameter to depths ranging from 50 to 200 feet. Post-shell excavation of a trench through the crater along this diameter revealed the situation shown in Figure 3. The true crater and rupture zones were fairly well defined by these columns. Of particular interest were the final locations of columns 9 and 13, which were extended and folded back over the edge of the true crater. The other columns show very strong effects of shear and rupture. From these data the depth of the true crater is believed to be 128 feet and the radius 150 feet. The depth of the rupture

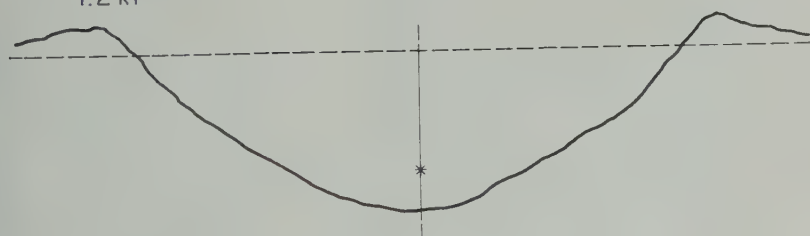
JANGLE S
1.2 kt



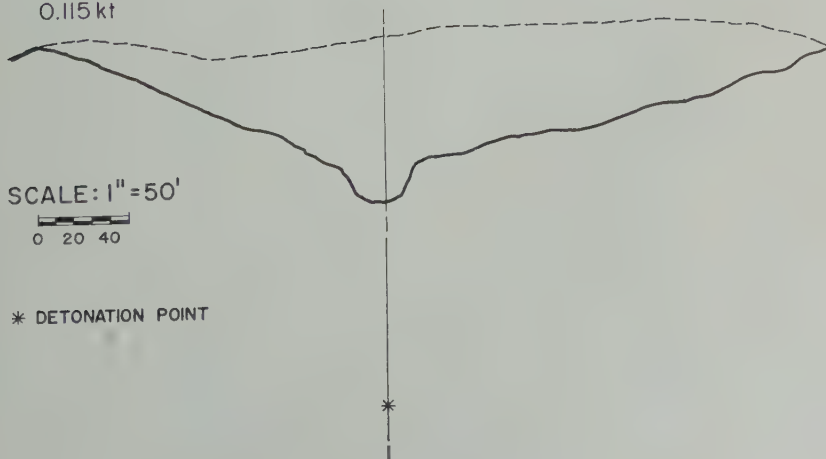
JANGLE U
1.2 kt



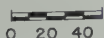
TEAPOT
1.2 kt



NEPTUNE
0.115 kt

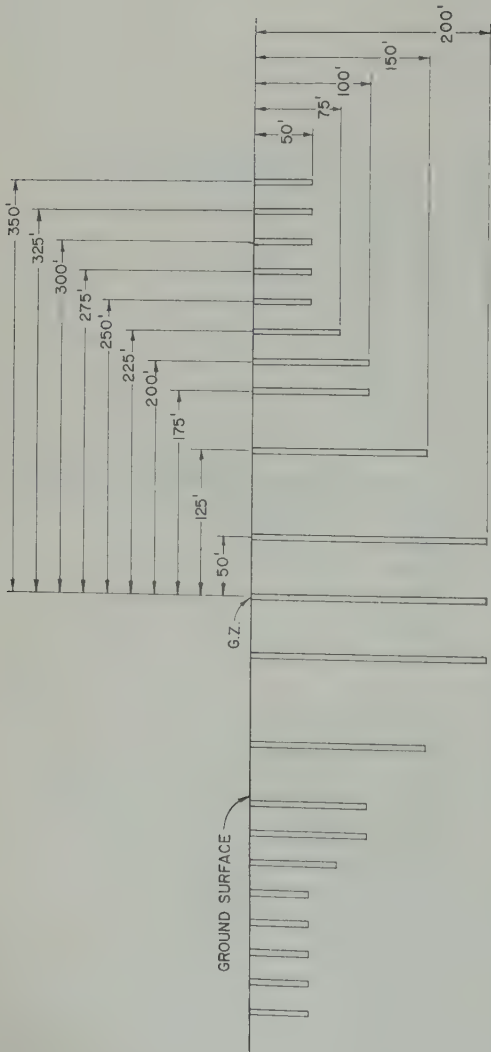


SCALE: 1"=50'

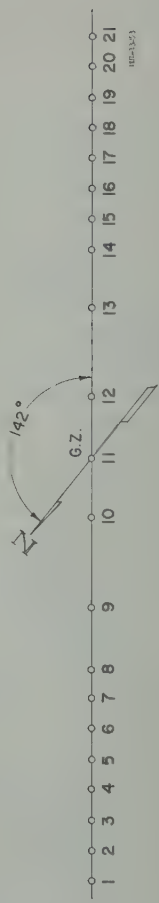


* DETONATION POINT

Fig. 1. Nuclear crater profiles.



VERTICAL CROSS SECTION



PLAN VIEW

Fig. 2. Pre-shot locations of sand columns for Teapot Ess.

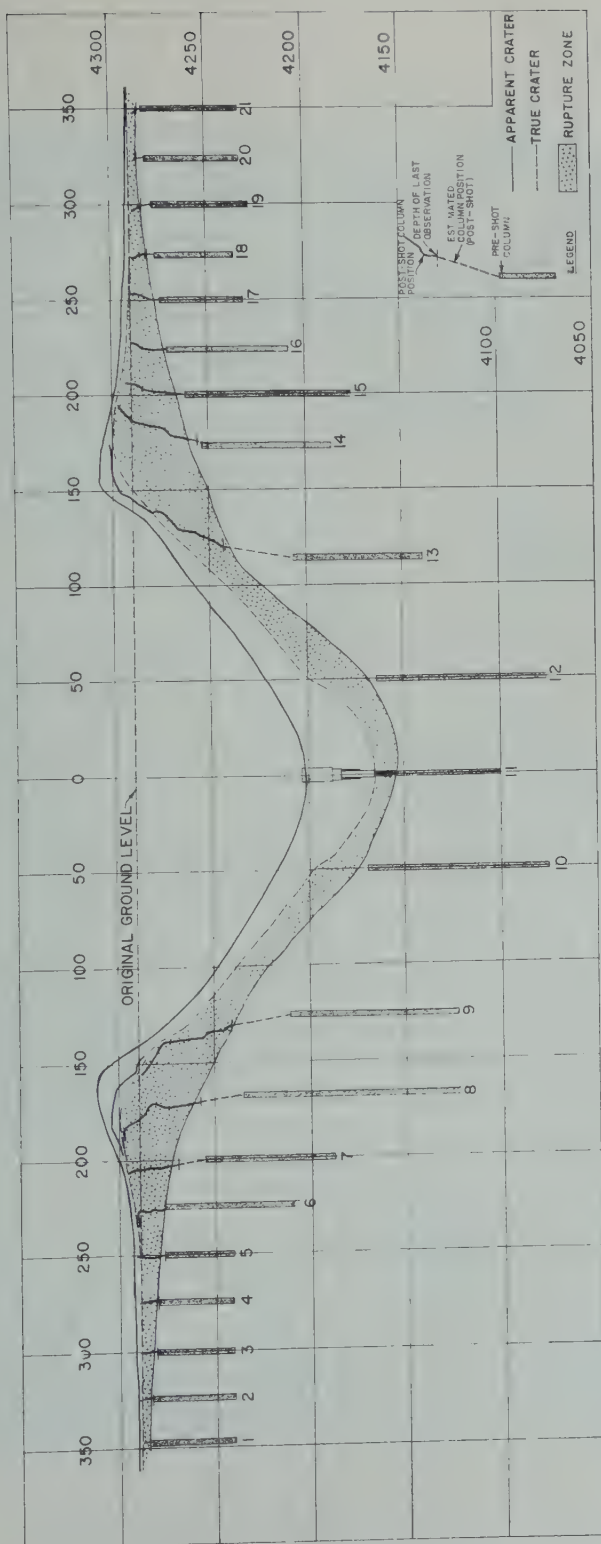


Fig. 3. Post-shot locations of sand columns for Teapot Ess.

zone can only be estimated, but its radius is believed to be 250 to 275 feet.

Figure 4 is an aerial view of area 10 showing Teapot Ess and Jangle U in top center. The trench in Teapot Ess is clearly visible. The recent high-explosive alluvium craters described by *Murphey and Vortman* [1961], including Scooter in the lower center and the three Stagecoach craters on the left, are also visible.

Neptune. The Neptune event occurred during Operation Hardtack, phase II, on October 14, 1955 [*Shelton, Nordyke, and Goeckermann*, 1960]. A 115-ton nuclear device was fired at a point 100 feet below a 30° slope in bedded tuff. The vertical distance to the surface was 110 feet. The zero-point room was 12 by 17 by 10 feet high with a concrete floor. The tunnel configuration was a buttonhook in shape and was stemmed before the shot in several places with sandbag plugs.

Upon detonation, the surface rose in a hemispherical dome to a height of 25 to 35 feet before it was disrupted by ejected material and venting gas, with large rocks going 80 to 100 feet in the air. A large dust plume was formed which rose to a height of about 1000 feet. A large mass of rock and debris cascaded down the slope (upwind), carrying small amounts of radioactivity into gullies as far away as 2000 feet.

The shape of the crater formed by the Neptune detonation was influenced by the slope of the surface in that almost all the debris formed a slide originating at the lower edge of the crater and terminating about 800 feet down the slope. The mean diameter was 200 feet, and maximum depth was 35 feet.

Eleven holes have been drilled into the region surrounding the Neptune detonation to determine the physical state of the rock and to delineate the radioactive regions [*Thompson and Misz*, 1959]. Figure 5 shows many of these holes and the post-shot state of the medium derived from them. Most of the layers overlying the shot retained their continuity but collapsed into the cavity produced by the explosion. The mixing that occurred was minor, and the different lithologic units are still easily identifiable.

Crushing of the tuff occurred to a distance of 40 feet downward and 50 feet laterally, except in the direction of the original drift where crushing extended to 80 feet. The extent of crushing was apparently influenced by bedding-

plane weaknesses. Fracturing of the material extended to 70 feet in the hemisphere below zero point, according to interpretations of core logs. Above the original zero point, fracturing extended to the surface, the boundary of fracturing lying on a cone whose base extended slightly beyond the surface crater region.

Integration of the total fallout patterns and the surface indicates that 0.5 per cent of the total fission product activity produced by the explosion escaped from the crater. Certain volatile isotopes present at early times of the explosion were enriched by a factor of 10 for Sr^{89} , Sr^{90} , and Cs^{137} .

Discussion. The results of these four nuclear explosion craters are shown in Figures 6 and 7 along with all the pertinent high-explosive data for alluvium [*Sachs and Swift*, 1955; *Doll and Salmon*, 1951; *Lewis*, 1958; *Murphey*, private communication, 1961]. Both nuclear and high-explosive data have been plotted, using Wilkins scaling. This type of scaling has been derived on the basis of high-explosive data alone and outlined by *Vaile* [1961] and also by other investigators [*Chabai*, 1959; *Pokrovskii and Fedorov*, 1957].

As can be seen from the plots, the nuclear craters (with the exception of Jangle S) fall well within the scatter of the high-explosive data. Hence it can be concluded that, for subsurface detonations, the craters produced by nuclear explosions are quite comparable with those produced by equal yield chemical explosions.

Interestingly, the Neptune point falls on the alluvium curve even though it was in tuff. The effect of the hillside may have just compensated for the decrease in crater dimensions expected for a cratering detonation in a soft rock and it produced a crater with dimensions proper for alluvium.

MECHANISMS OF EXPLOSIVE CRATER FORMATION

The general subject of cratering has been studied for many years by many investigators. Much of this work has been of a qualitative nature, with a few notable successes for quantitative analysis of isolated phases of the process involved. Much of the analysis has been based upon empirical relationships or dimensional analysis arguments which, though providing a useful bridge, have not given much insight into the basic problems involved. I shall set down










Fig. 4. Aerial view of area 10 showing craters.

U12C-03-NEPTUNE CROSS SECTION B-B'









THOMAS L THOMPSON, Geologist - S R I.
JOHN B MISZ, Geologist - L R L - N
AUGUST, 1959

LEGEND

LITHOGRAPHY & ENGINEERING MATERIALS

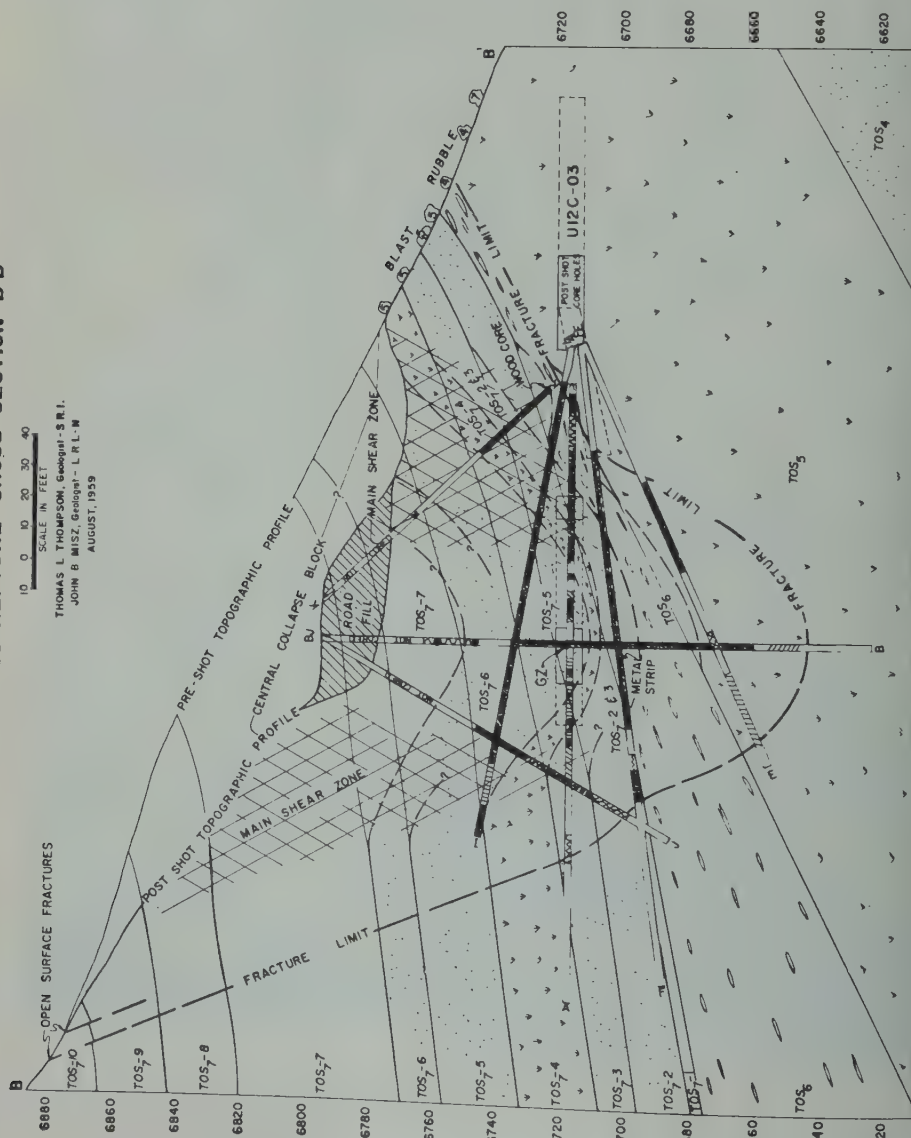
- | | | | |
|---|----------------------|---|---------------------------------|
|  | WELDED TUFF | } | Hard, Brittle Rock |
|  | PORCELLANITE | | |
|  | GRANULAR TUFF | } | Hard to Friable Rock. |
|  | SANDSTONE | | |
|  | FRIABLE SANDY TUFF | } | Structurally Weak Rock |
|  | LAPILLI TUFF | | |
|  | 10% LITHIC FRAGMENTS | } | Light, Porous Compressible Rock |
| | | | |

CORE HOLE ANALYSIS

-  SOLID CORE
 LONGITUDINAL FRACTURES
 BROKEN CORE
 MUD AND SMALL FRAGMENTS
 POOR RECOVERY
 DRILLING FLUID LOSS
 WOOD FRAGMENTS
 HIGH RADIATION

OTHER SYMBOLS

- 5 SURFACE RUBBLE SHOWING
SUBUNIT NUMBER
JO DRILL COLLAR COORDINATES
TO TOTAL DEPTH
— PRE-SHOT LITHOLOGIC CONTACT
— DISPLACED LITHOLOGIC CONTACT
MAIN SHEAR
ROAD FILL



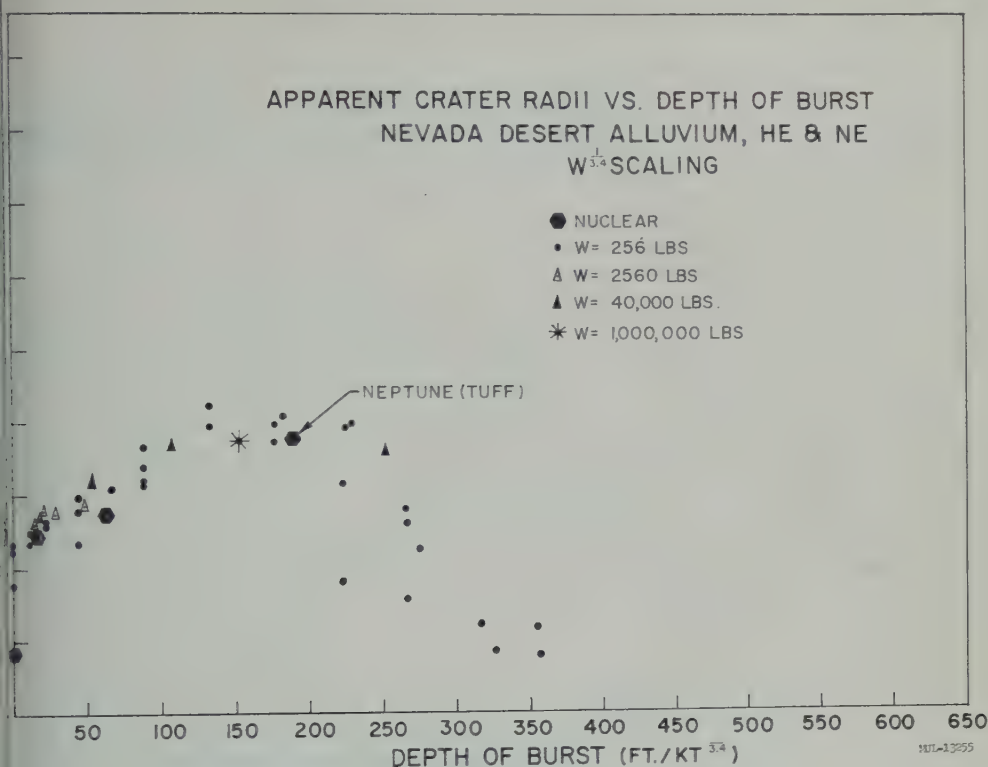


Fig. 6. Scaled crater radius vs. depth of burst, desert alluvium.

is relating to the mechanisms involved in the formation of explosive craters, and give estimates of their relative importance. The analogy between explosive craters and impact craters, such as result from a meteor hitting Earth or the moon, is not complete. I shall attempt to point out some of the differences.

Ideas expressed in this paper are based on data from various sources, ranging from the experimental results derived from cratering programs such as the Sandia, Scooter [Murphey and Vortman, 1957], and Neptune re-entry and recon-struction [Shelton, Nordyke, and Goeckermann, 1961], to theoretical advances made possible by the development of machine calculations using the UNEC code described by Maenchen and others [1961].

For the purpose of this paper, the definitions of the terms to be used should be defined. Unfortunately there is no uniform terminology in the field of elastoplastic behavior, and each investigator is forced to define his own terms for his

purposes. This difficulty, of course, arises to a great extent because of the tremendous range of properties of materials, which causes definitions that are adequate for one material to be unsuitable for another.

Figure 8 is a schematic drawing of a typical crater cross section showing the pertinent parameters. The apparent crater is defined as the crater that is visible on the surface; its dimensions are measured with respect to the original ground level. The true crater is defined as the boundary between the loose, broken fallback material and the underlying material that has been crushed and fractured but has not experienced significant vertical displacement. The products of the explosion are widely dispersed throughout the fallback material that rests in the true crater after the explosion.

The rupture zone is perhaps the most difficult to define, particularly with regard to differentiating it from the plastic zone. There is, of course, a gradual transition from one zone to the other. In the rupture zone near the true crater inter-

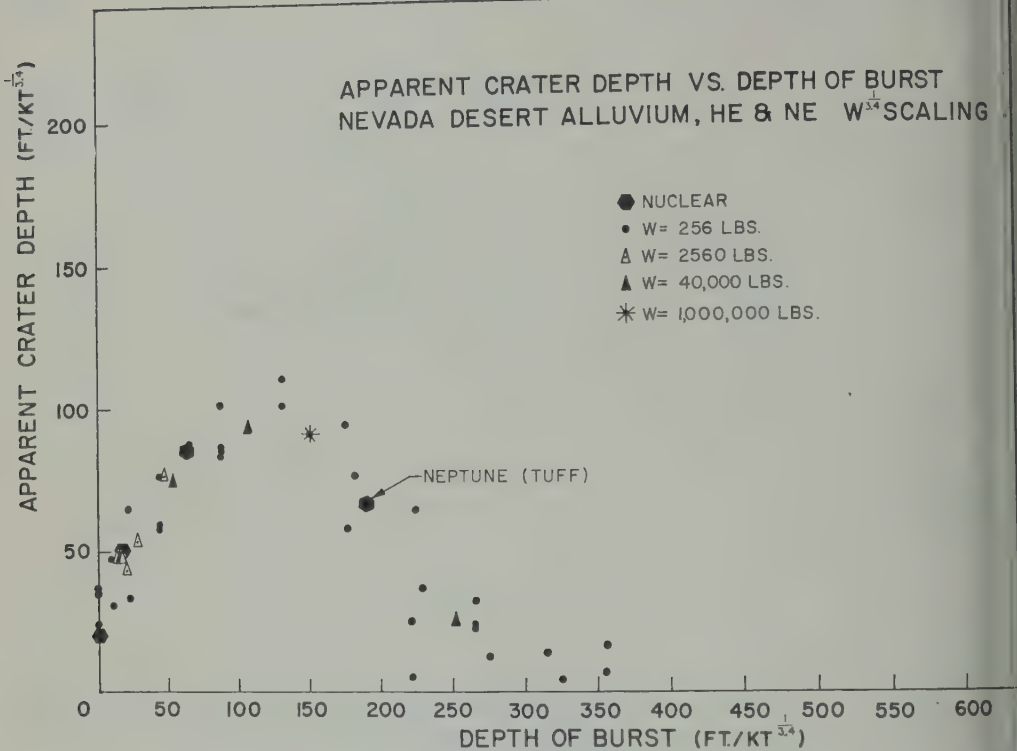


Fig. 7. Scaled crater depth vs. depth of burst, desert alluvium.

face, large amounts of fracturing and crushing by shear failure and gross displacements by faulting and underthrusting are generally seen. Their severity decreases with depth into the rupture zone, until near the rupture zone/plastic zone interface only small-scale shear failures (of the order of an inch) are found. This zone gradually shades into the plastic zone, in which there are small uniform permanent displacements which decrease to infinitesimal values as one goes into the elastic zone.

The extent of these zones is very dependent on medium, varying widely as one goes from soft medium like alluvium to a hard medium like basalt. The definitions given here have been derived principally for soft materials such as alluvium. For a discussion of cratering in hard rock some of these terms would have to be redefined.

Mechanisms of crater formation. One phenomenon present to varying degrees in underground explosions is the crushing, compaction, and plastic deformation of the medium immediately surrounding the source of explosion, whether it be a chemical, nuclear, or impact explosion. As the high-pressure gases generated by the explosion push on the walls of the cavity, a shock wave is generated across whose spherical surface there is a sharp discontinuity in the physical state of the material. This discontinuity propagates outward at a velocity that, for high pressures, is faster than the speed of sound in the medium.

For chemical explosives the initial pressures are of the order of 100 to 200 thousand atmospheres.

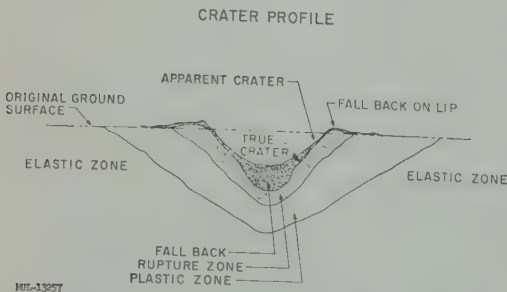


Fig. 8. Schematic drawing of a typical crater.

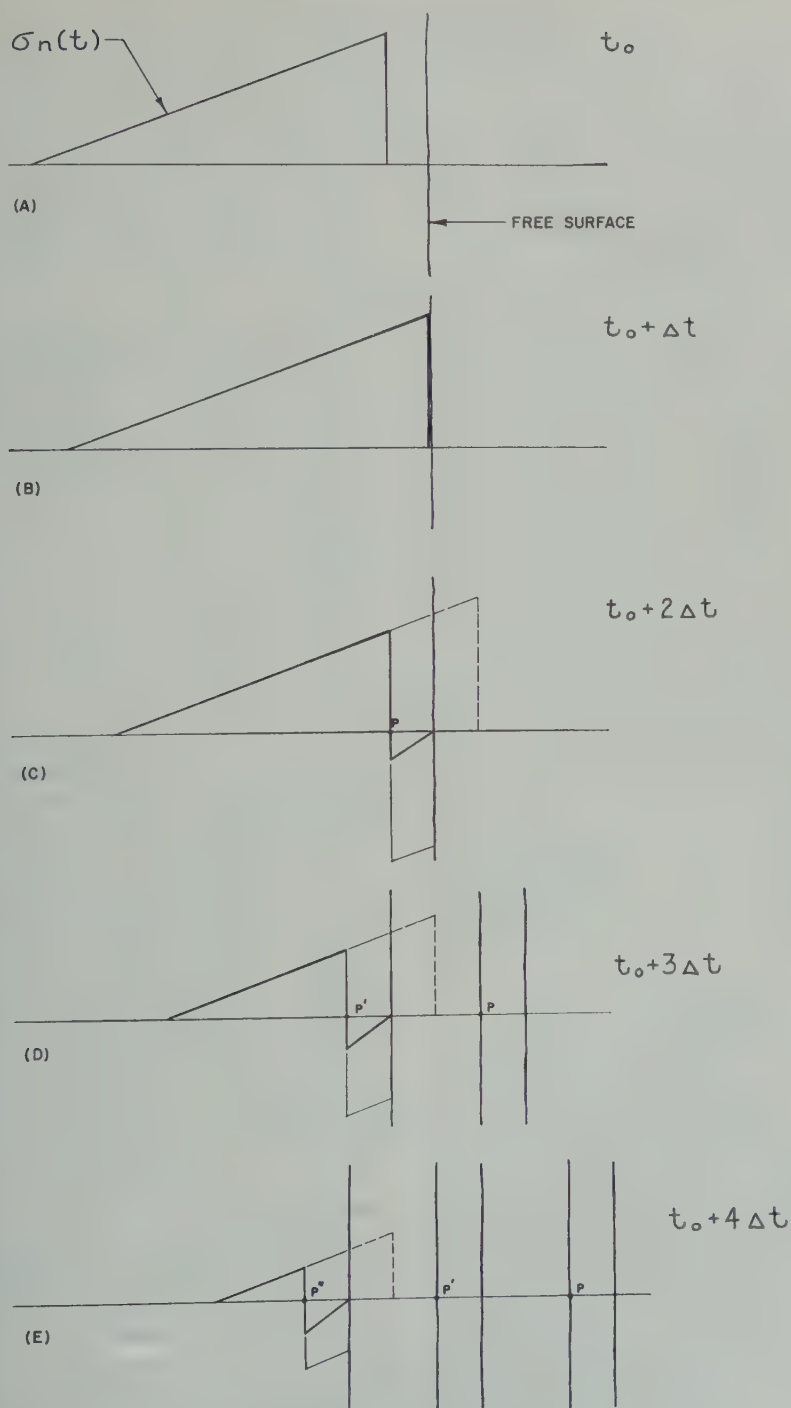


Fig. 9. Schematic drawing illustrating the spall mechanism.

pheres; for a nuclear explosive they are as large as 10 to 100 million atmospheres, depending on the initial cavity; and for a meteoritic impact explosion they can have any value ranging between these two, depending on the meteorite's velocity and the type of material it hits. For explosions where the pressures are greater than 500 thousand atmospheres, the medium is melted and vaporized when the shock passes through it. As the shock wave moves outward in a spherically diverging shell, the peak pressure in the shock front drops because of spherical divergence as well as energy expenditure in doing work on the medium. For pressures above the dynamic crushing strength of the material this work appears in the form of crushing, heating, and physical displacement. In regions outside the limit of crushing, the shock wave will still produce permanent deformation by plastic flow until the peak pressure in the shock front has decreased to a value equal to the plastic limit for the medium. The plastic limit marks the boundary between the elastic and plastic zones described for Figure 8. As with the definitions given for Figure 8, the limits of crushing and plastic deformation vary widely from material to material.

The above picture of the first few milliseconds of an explosion neglects the effects of any free surface—effects that are exceedingly important. As a compressive wave encounters a free surface, it must match the boundary condition that the pressure, or more correctly the normal stress, be zero at all times. This results in the generation of a negative stress wave or rarefaction which propagates back into the medium. This process is shown schematically in Figure 9, where for simplicity a triangular-shaped stress wave, $\sigma_n(t)$, has been assumed instead of the more correct exponential shape. At some depth, such as P in Figure 9, the sum of the two stress waves equals the dynamic tensile strength of the medium. The medium breaks in tension at P , and a piece flies off with a velocity characteristic of the total momentum trapped in the piece. This produces a new free surface that will break at P' and again at P'' . For a loose material like alluvium, this process (called 'spall') makes almost every particle fly into the air individually, whereas in a rock like basalt the thickness of the slabs is generally determined by the presence of pre-existing joints and zones of weakness. For the

case of a small sample, or where there is a massive block, the dynamic tensile strength of the rock determines the thickness of spall. As the distance from the explosive to the free surface increases, the peak pressure decreases, so the maximum possible tensile stress decreases until it no longer exceeds the tensile strength of the medium. In addition, the velocity given to the spall decreases in proportion to the pressure.

For ranges beyond the point where spalling occurs, the negative stress in the rarefaction wave will decrease the shear strength of the medium, which results in large plastic deformations and ruptures. This makes the ruptures extend a considerable distance along the surface and contributes to the formation of the crater. Ultimately, the surface expression of an underground explosion is only a small excursion of the surface. Spalling of the surface is probably the most important phenomenon in cratering, especially for shallow depth of burst, and is the easiest mechanism to observe and to calculate.

The third mechanism of importance in cratering, particularly for craters deeper than a critical depth, is what I have termed 'pressure acceleration.' This is a long-period acceleration given the material above the explosion by the adiabatic expansion of the gases trapped in the cavity. For some cases, particularly for shallow depths of burst, this gas also gives appreciable acceleration during its escape through cracks extending from the cavity to the surface. At very shallow depths of burst the spall velocities are so high that the gases are unable to escape any pressure before venting occurs. For deep explosions, the acceleration given to the overlying material is so small as to be negligible. This process will be examined in more detail in a later section.

Subsidence is the last remaining major process that makes a significant contribution to the formation of the apparent crater. It is very closely linked to the first process of compaction and plastic deformation, without which there would be no void into which material could subside. Subsidence occurs when the spall or the acceleration has so distended the overlying material that large cracks are produced through which the explosion gases escape. This material having been fractured and crushed by the shock

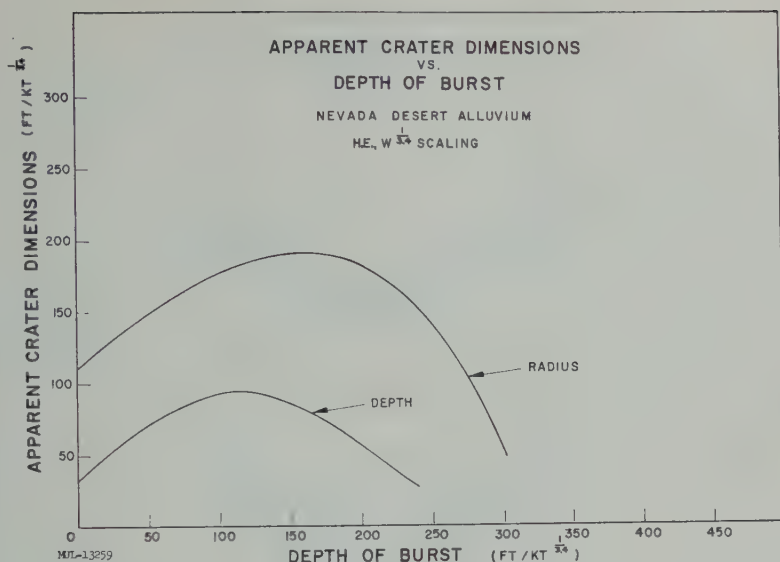


Fig. 10. Apparent crater dimensions vs. depth of burst.

collapses into the cavity. Subsidence is important, of course, for very deep explosions.

Effect of depth of burst. The part each of the mechanisms plays in producing a crater is strongly dependent on the scaled depth of the explosion. The effect of depth of a 1-kiloton explosion (equal to 10^{12} cal) on the radius and depth of the crater is shown in Figure 10. These are based on the chemical explosive data in Figures 6 and 7, using $W^{1/3.4}$ scaling. Explosions in other media would give different dimensions; for example, dimensions would be as much as 20 to 30 per cent smaller for rock (see *Armstrong and Company, 1948 a and b*) and 20 per cent larger for water-saturated media (see *Armstrong and Company, 1948c*).

The data for alluvium, along with much other information, have been used for the sketches in Figure 11 of four typical craters in cross section. The small dashed circle about the detonation point indicates the size of the original TNT

Surface burial. As can be seen in the top sketch in Figure 11, the crater resulting from the burial of an explosive very near the surface of the ground is produced to a large degree by the spalling action and plastic deformation. There is also a significant action by the gases in the initial gas

sphere which erodes the surface of the crater, but it is not a significant mechanism in the formation of the crater. The radius is extended to its limit by spalling action resulting from a horizontally diverging shock wave, but the major process for the depth of the crater and for lip formation is the plastic deformation and flow of the material in the rupture zone. Very little fallback is found in a crater of this kind, and the true crater and apparent crater are almost the same.

This sketch is based on data from the nuclear crater Jangle S, which has dimensions somewhat smaller than would be expected from a chemical explosion at the same depth of burst because a large fraction of the energy from a nuclear explosion is released in the form of thermal and X-ray radiation. A meteoritic impact explosion could be considered analogous to a surface burial only for very low-velocity meteors, in which case the pressures involved would be more like those occurring in a chemical explosion than in a nuclear explosion, so that dimensions more like those shown in Figure 10 for a surface burst would be expected. The mechanisms involved, however, would be essentially the same as for the surface nuclear explosion.

Shallow burial. A cross section of the crater resulting from shallow burial of the explosive is shown in the second sketch in Figure 11. This

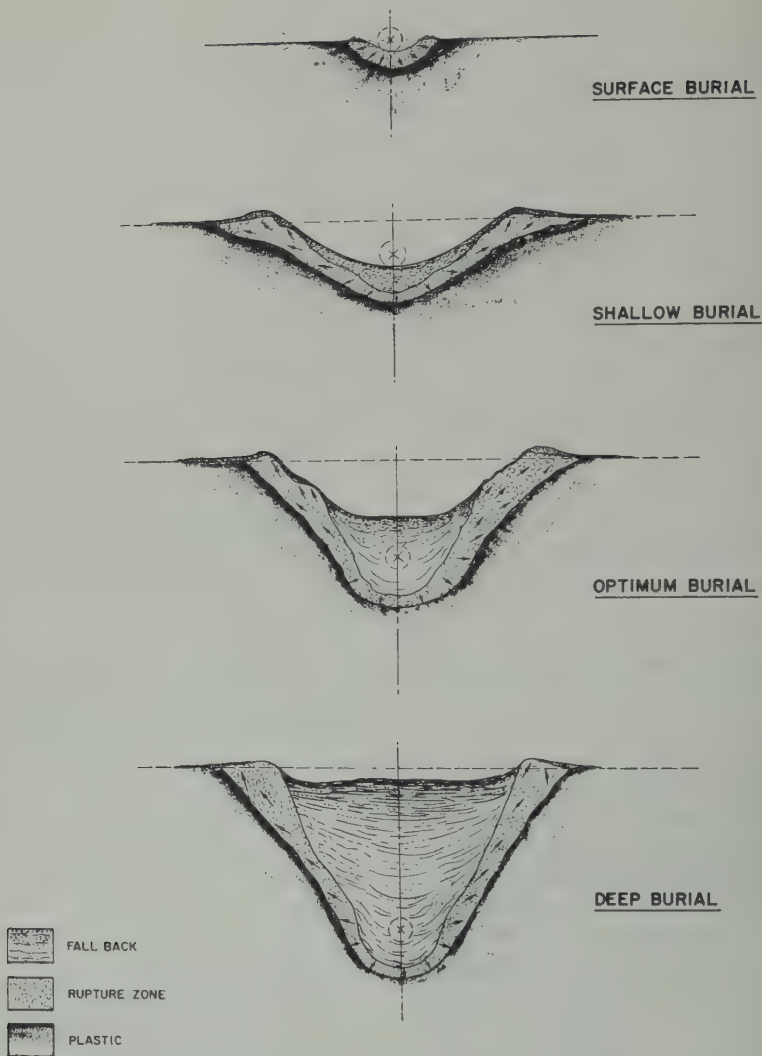


Fig. 11. Typical crater profiles vs. depth of burst for alluvium.

scaled depth of burst corresponds roughly to the scaled depth of the nuclear explosion Teapot Ess (66 feet). Spalling of the free surface has now become the dominant process for the formation of the crater. Gas acceleration and scouring action are of only minor importance because of the high velocities given to the material by the spalling process. The radius of the crater is determined by the limit of the spalling process, whose velocities decrease rather rapidly with increasing surface radius. This decrease of spall velocity with radius leads to the 'folding back' of the material on the edge of the crater to form

the lip that is evident in many craters, including Teapot Ess and Meteor Crater, Arizona [S. J. Singer, 1959]. Below the original surface, the radius of spall decreases because of the increased total path length which the shock and the material must travel. This results in a roughly parabolic-shaped true crater. The extent of fallback and rupture zone for a crater from a shallow depth of burst were very well defined by the sand column techniques used on Teapot Ess and the sketch is based very closely on this work.

The impact and explosion of a high-velocity meteor is probably most closely simulated

explosion at a shallow depth of burial. For example, Shoemaker's analysis shows that the Scooter crater corresponds to an explosive-buried crater with scaled depth of burst of 145 feet [Shoemaker, 1959]. Further, most general features noted in meteor craters and in explosion craters for shallow scaled depth of burst. A nuclear explosion would more closely simulate the meteor impact explosion than a chemical explosion would, because of the high initial pressures and energy densities in both nuclear and meteor explosions. Meteor explosions are characterized by relatively small amounts of condensable explosion products and vaporization of considerable quantities of medium surrounding the explosion. However, experience has shown that, for shallowly buried explosions in alluvium, there are very small differences between chemical and nuclear explosion craters, presumably because (1) gas radiation is not important for shallow depth of burst, and (2) there is 10 to 20 per cent water in alluvium, which produces noncondensable gas in the cavity when vaporized. Condensable gases, i.e., gases that condense at a relatively low temperature, such as silica vapor, drop out of the vapor phase relatively early in the explosion and do not contribute to the gas pressure in the cavity. Water vapor and carbon dioxide do not condense and hence add to the gas pressure.

The Scooter crater is a virtually complete venting of the radioactive material from the Teapot Ess explosion. This leads to the conclusion that, if the high pressures and temperatures predicted by Shoemaker are correct, the meteoritic material from a meteor impact would also be completely vaporized and blown out to the atmosphere and be spread over a wide surrounding countryside. It should be added that there is undoubtedly not an exact correlation between an impact crater and an explosion crater because the meteor's energy is released in the form of a line source as opposed to a point source for an explosion. Thus, deviation from dimensions predicted for explosion craters should be expected.

Optimum burial. For an explosion at optimum depth of burial, i.e., at a depth that gives maximum apparent crater dimensions, the resulting crater would appear as shown in the third sketch in Figure 11. The apparent dimensions shown are taken from the

Scooter crater [Murphy and Vortman, 1961]. The Scooter event was the detonation of a 1-million-pound TNT sphere buried 125 feet deep in desert alluvium, a scaled depth of burst of 153 feet. (The scaled depth is the corresponding depth for a 1-kiloton explosion.) The true crater and rupture zone both for this sketch and for the deep burial (bottom sketch) are only estimates; there were no post-shot excavations. For craters at these depths of burst, all three phenomena—plastic deformation, spall, and gas acceleration—are important, but the last has become the dominant feature of the cratering process. When the shock wave reaches the surface, it has decayed to the point where, although it is still capable of fracturing the material in tension (since most media have very small tensile strengths), the velocities given the material are relatively small. Because the maximum height to which a particle will go is proportional to the square of the initial velocity, the throwout would not go any appreciable distance into the air if spalling were the only process.

The inadequacy of the spall mechanism and the necessity for some kind of gas acceleration is best seen from the surface-motion data from Scooter obtained from high-speed motion pictures of seven surface targets [Feigenbaum and Wegkamp, 1961]. Taking the displacement-vs.-time data, and computing vertical velocities, we get the plots shown in Figure 12. The straight lines are least-squares fits to the data over the ranges of 80 to 350 m/sec and above 350 m/sec. General venting which obscured the targets occurred at about 1.2 sec. As can be seen from Figure 12, the data can very easily be broken down into these two regions, one with a negative acceleration and the other with an approximately uniform positive acceleration. The data from the graphs in Figure 12 are summarized in Table 2 where the initial velocities and the velocities at 1.2 sec are given, along with the accelerations derived from the slope of the lines in Figure 12. In addition, the maximum heights to which a particle would go are computed for both the initial velocity and the 1.2-sec velocity, using the equation $h = v^2/2g$. Heights from the 1.2-sec velocities are quite consistent with observed particle trajectories, indicating that there was no appreciable acceleration after venting occurred.

The picture presented by these data is fairly

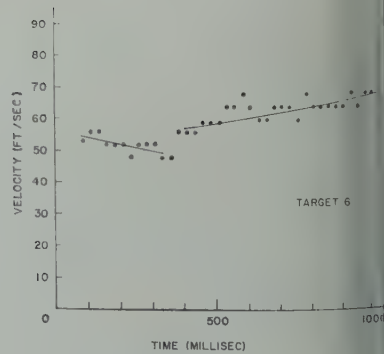
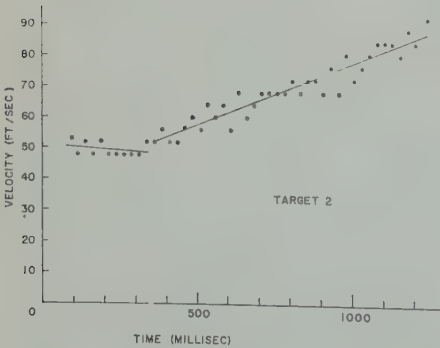
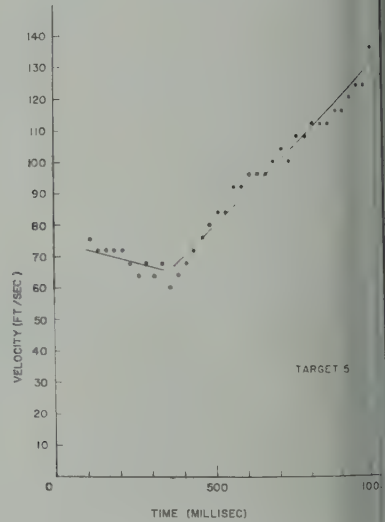
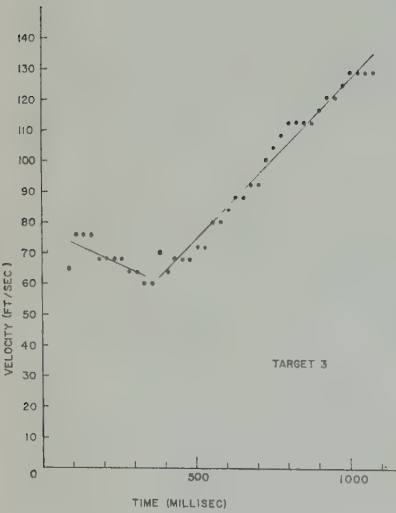
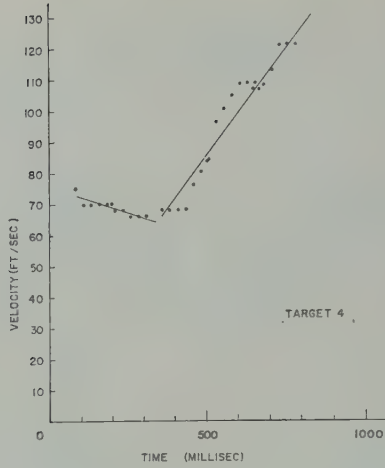


Fig. 12. Velocities of surface targets vs. time, Project Scooter.

E. 2. Summary of Acceleration, Velocity, Displacement Data for Scooter Surface Targets

Numbers	2	3	4	5	6
Distance from surface zero of burst, ft	60	30	0	30	60
Initial velocity, ft/sec	50	74	73	73	54
Acceleration 80-350 m/sec, ft/sec ²	-8	-46	-34	-27	-20
Velocity at 1.2 sec, m/sec, ft/sec ²	+39	+102	+133	+100	+17
Shock wave height at initial velocity, ft	85	145	177	150	70
Shock wave height at 1.2 sec, ft	106	328	490	352	77

The first motion experienced by the surface ground is that produced by spall. The spall velocities of about 70 to 75 ft/sec are realized at surface zero, decreasing with distance from this point because of the increased length for the shock wave. As the rarefaction propagates back toward the cavity, all material is given an upward velocity which increases in magnitude with depth. After the passage of the shock, the material is in approximately free fall, as shown by the magnitude of negative acceleration in the period from 80 to 350 m/sec. When the rarefaction reaches the cavity, the cavity begins to expand very rapidly, pushing up the loose and broken layers above it, giving up each layer as it moves upward, giving them all to the same velocity in much the same manner that an engine of a freight train accelerates its cars when it reverses direction. Ultimately this second push reaches the surface and all the earth or rock above the surface is moving as one mass. In Scooter, the hemispherical surface set in motion by spall experienced a relatively uniform positive acceleration starting at about 350 to 400 m/sec. The rate of acceleration was very dependent on radial distance from surface zero. Finally, the divergence of the hemispherical plug causes large cracks to open from the cavity to the surface, through which the high-pressure gas escapes. During its escape, it gives appreciable

acceleration to the material through which it is passing. As a result the surface layers experience a much longer period of acceleration than the deep layers. Much of the material immediately above the cavity does not attain the high velocity of the surface and falls back in place with very little mixing or disruption of the stratigraphy.

An order-of-magnitude estimate of the validity of this picture can be made, based on the results of a UNEC calculation [Maenchen and Nuckolls, 1961] and some simple concepts. The UNEC code is a program for the IBM 7090 which can make a one-dimensional elastic-plastic-hydrodynamic calculation of the early history of an underground explosion. Calculations for Scooter give the results that, at the time when the shock wave reaches the surface of the ground, the pressure of the gas in the cavity is approximately 175 bars. The cavity radius is about 36 feet, vs. an initial radius of about 15 feet for the TNT sphere. The initial surface velocity predicted by UNEC is 103 ft/sec, a number in fair agreement with the observed values of 70 to 75 ft/sec. This agreement is particularly encouraging when the difficulties of making an elastic-plastic calculation for a sand-gravel mixture are considered.

With these numbers, an estimate can be made of the magnitude of the gas acceleration by considering the material contained in the solid angle, Ω , above the cavity. This solid angle is defined by the approximately hemispherical surface which is the first evidence of surface motion. Figure 13a shows a sketch of the situation at 350 m/sec if the material above the cavity continues to move but the cavity remains at the 36-foot radius. Numerous voids are opened up whose total volume equals the volume of the hemispherical segment. Figure 13b shows the configuration if the cavity is allowed to expand to take up all these voids. The new cavity volume is now $V_0 + \Delta V$, where ΔV is the volume of the spherical segment. Thus

$$\Delta V = [\pi h^2(3R - h)]/3$$

where R = radius of hemisphere = depth of burst + h , and h = height of hemispherical segment. The new pressure in the cavity after this adiabatic expansion is given by

$$P = P_0 [V_0 / (V_0 + \Delta V)]^\gamma$$

where P_0 = original pressure in cavity.

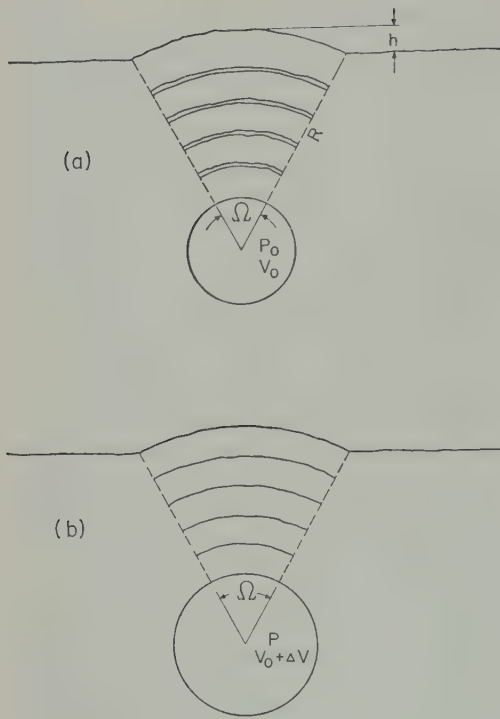


Fig. 13. Schematic drawing of optimum depth of burial for (a) no expansion of cavity, (b) cavity expanded to take up voids.

If we now assume that this conical mass moves as one under the influence of the gas in the cavity, we have

$$PA = Ma$$

where

M = mass of cone $\cong \Omega R^3 \rho / 3$.

A = area of the truncated cone

$$= \Omega \left(\frac{V_0 + \Delta V}{4\pi/3} \right)^{2/3}.$$

a = acceleration experienced by conical mass.

Thus we have

$$a = \frac{\Omega P_0 \left(\frac{V_0}{V_0 + \Delta V} \right)^\gamma \left(\frac{V_0 + \Delta V}{4\pi/3} \right)^{2/3}}{\frac{R^3 \rho \Omega}{3}}$$
$$= 1.15 \frac{P_0 (V_0 + \Delta V)^{2/3}}{R^3 \rho} \left(\frac{V_0}{V_0 + \Delta V} \right)^\gamma$$

At 350 m/sec the observed height of the spherical segment for Scooter was approximately 16 feet

(490 cm). Using 125 feet (3810 cm) for the depth of burst, $R = 4300$ cm and

$$V_0 = 5.55 \times 10^9 \text{ cm}^3 \quad \Delta V = 3.12 \times 10^9$$

The density of alluvium is about 1.6 g/cm³. The γ for TNT at 175 bars as obtained from Jones and Miller [1948] is about 1.3. Using $P_0 = 175$ bars = 1.75×10^8 dynes/cm², we have

$$a = 1.15$$

$$\times \frac{1.75 \times 10^8 (8.67 \times 10^9)^{2/3}}{1.6 (4.3 \times 10^3)^3} \left(\frac{5.55}{8.67} \right)^{1.3}$$

$$= 6.68 \times 10^3 (0.64)^{1.3}$$

$$= 3.74 \times 10^3 \text{ cm/sec}^2$$

$$= 123 \text{ ft/sec}^2$$

Comparison of this number with the accelerations observed in Table 2 for the period from 350 to 1200 m/sec shows that it is remarkably close to those observed, particularly for central targets. It is considerably higher than the observed accelerations of the targets 60 ft from the surface zero, even when one considers that the numbers given in Table 2 are vertical components. One would not, of course, expect the acceleration to be constant in time, but to decrease with time because of the dropping pressures in the cavity. However, the continuing acceleration of the surface layers by the escaping through cracks and fissures gives additional acceleration to compensate for the drop. Obviously, such a picture is inadequate but it does indicate the correct order of magnitude of these effects.

The sequence and magnitude of events described above apply only for a medium of alluvium. For other media the numbers, and consequently the relative importance of the various mechanisms, are greatly different. Preliminary results from high-speed motion pictures of ground motion at surface zero for the three large Buckboard detonations [Murphy and Vortman, 1961] provided some interesting information for another medium. This cratering program in basalt, conducted at the Nevada test site, included three 40,000-pound detonations: shots 11, 12, and 13, buried at depths of 25.5, 42.7, and 58.8 feet, respectively. These correspond to scaled depths of burst of 80, 140, and 186 feet, using $W^{1/3.4}$ scaling to 1 kilo

of comparison with the alluvium data. Surface-motion data for these three shots are in Table 3 along with the spall velocities indicated by the UNEC code.

agreement here is excellent. The higher velocities are attributable to the much greater strength and competence of the basalt. A shock is attenuated much less in traversing a foot of basalt than a foot of alluvium. The result is that, for a scaled depth of burst equal to the optimum's, the spall velocities are more than as large in basalt as in alluvium. This indicates that the mechanics of cratering for an optimum depth of burial in basalt are much like those for a shallow depth of burial in alluvium. The gas acceleration does not have time to accelerate the rock, and the crater is formed almost entirely by spalling. For deep burial in basalt the spall velocity is not so high, the mass of material to be accelerated by the explosion is so large that the acceleration would be very small.

Considering impact cratering mechanisms, it is obvious that gas acceleration must play a less significant role, since the hole produced by the entry and explosion of the meteor would permit entrapment of the gases produced by the explosion. The small amount of work that has been done on the effects of the stemming of a cratering charge [Martin and Hinze, 1958], for relatively shallow depths of burst, indicates that apparent crater dimensions are reduced by stemming like 15 to 30 per cent by eliminating stemming. There are no data on the effect of stemming at depths of burst near optimum. The effect of gas acceleration would be more important, although this work does show that the effect of stemming increases with depth of burst.

Deep burial. The bottom sketch in Figure 11 shows the cross section of a crater resulting from the detonation of an explosive at a depth beyond the optimum depth of burial. The present crater profile here follows closely the best Stagecoach crater [Murphey and Vortman, 1961], whose scaled depth was 253 feet. Spall velocities are now very small, about 15 ft/sec. The direct gas acceleration is at best an order of magnitude smaller than for optimum depth and is in general difficult to quantify. The acceleration resulting from frictional drag by the escaping gases is probably

most important for this situation. The fallback within the crater should be well ordered with little or no disruption of the stratigraphy as indicated in the sketch. The lips are produced, to a very large extent, by the rupture and plastic flow of the material in the rupture zone. About 50 to 75 per cent of the apparent crater volume can be attributed to subsidence, depending on the depth of burst.

Summary. In an attempt to give an over-all picture of the effects of these four mechanisms and their relative importance at various depths of burst, I have constructed a schematic diagram showing their effect on one crater dimension. For the purpose of this example I have chosen to use apparent crater depth because the role played by each mechanism is comparatively simple. Figure 14, a dimensionless graph, shows these relationships. The contribution from compaction and plastic deformation and subsequent subsidence is maximum for a surface detonation and decreases somewhat asymptotically with depth because of the increasing overburden pressure. The contribution of spall to apparent crater depth is, of course, directly proportional to the depth of burst for shallow depths of burst; it peaks and then decreases as shown for larger depths of burst because of the decrease in surface velocities. The effects of gas acceleration do not become significant until the spall phenomenon starts to lose its effectiveness. They

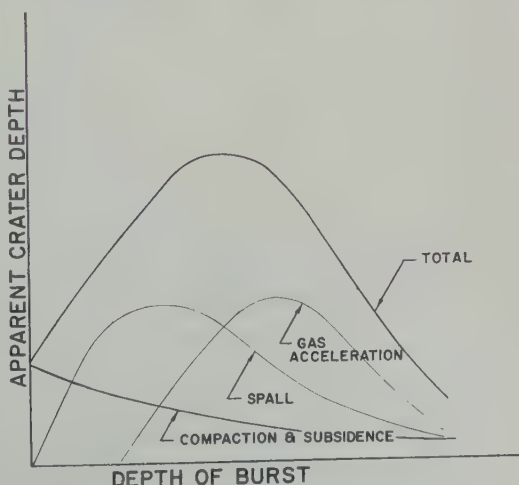


Fig. 14. Relative contributions of various mechanisms to apparent crater depth for explosion crater.

TABLE 3. Surface-Motion Data for Buckboard Shots 11, 12, and 13

Shot Number	11	12	13
Scaled depth of burst, ft*	80	135	186
Observed initial velocity, ft/sec	330	150	120
Calculated initial velocity, ft/sec	340	164	100
Residual cavity pressure, kb	10	9	8

* Scaled depths for these 40,000-lb (20 ton) shots are obtained by multiplying actual depths by the factor $(1000 \text{ tons}/20 \text{ tons})^{1/3.4}$, thus giving the corresponding depths for a 1-kiloton shot.

then increase somewhat to a peak, and tail off as shown. The effectiveness of even this tail is dramatically illustrated by films of the Blanca nuclear event [Johnson, Higgins, and Violet, 1959], an event somewhat similar in geometry to Neptune but with a scaled depth of burst about 50 per cent greater. The surface spall created no crater at all for this shot, but a large cavity or camouflet was produced that collapsed, with the subsidence progressing toward the surface. The collapse required 15 seconds to reach the surface, at which time the gases trapped in the cavity vented to the surface with a very startling plume of gas which rose about 600 feet in the air. When this venting first evidenced itself, some of the large surface rocks were ejected several hundred feet in the air by

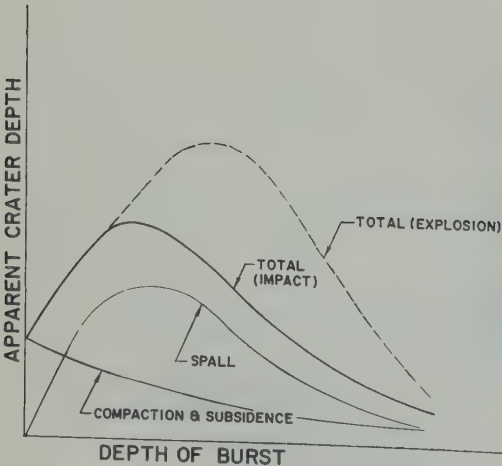


Fig. 15. Relative contributions of various mechanisms to apparent crater depth for impact crater.

the escaping gas, showing the particle-eject capabilities of these gases during their venting phase even for such a large scaled depth of burst. Summarizing these individual contributions, the crater depth gives the heavy line shown in Figure 14.

If we make the assumption that the principal difference between impact craters and explosion craters is the effect of gas acceleration, we shall eliminate this contribution and draw the picture shown in Figure 15. Thus, for surface and shallow depths of burst the craters should be quite similar, but for depths somewhat deeper than Teapot Ess, for example, serious differences would be expected to appear. However, if impact explosions are equivalent to surface and shallow-depth explosions, so that relatively good agreement with existing experimental data would be expected. For a set of conditions that would result in a deeper equivalent depth of burst than an impact explosion, the above-indicated deviations would be expected to occur.

A similar set of curves could be drawn for apparent crater radius, but it is very difficult to untangle the various effects. For surface detonations, spalling undoubtedly contributes to the radius to some extent. Gas acceleration is probably less important at all depths of burst in determining the radius than in determining the depth.

CONCLUSION

This discussion has been largely qualitative, but I believe the model outlined to be basically correct. Further development of this theory will require additional theoretical work as well as more experimental studies. The UNEC code has the limitation of being a one-dimensional code; it is valid only in the vertical direction and only until such time as the rarefaction arrives back at the cavity. At present, work is being done on a two-dimensional version of UNEC which will be much more useful for cratering purposes. This code will allow accurate calculation of the true crater and give initial spall velocities for all the material ejected into the air, and should adequately treat the early stages of the gas acceleration. Calculational treatment of the later stages, when venting occurs, is impossible with present codes. Attempts are being made to develop a treatment that will handle the problem. Further experimental work directed

exploring existing craters such as Scooter, geocoach craters, and the three large hard craters should be undertaken. Geomapping of craters has proved invaluable in analyzing the mechanics of cratering in the field and will be even more useful in the future with our experience with it.

We are on the edge of constructing a quantitative picture of cratering, and are hopeful that experimental and theoretical programs will permit the construction of a complete theory of the mechanism of explosive crater formation.

Acknowledgment. This work was done under the auspices of the U. S. Atomic Energy Commission.

REFERENCES

- A. J., Crater scaling laws for desert alluvium, *Armstrong Corp. Rept. SC-4391(RR)*, December 1959.
- B., and V. Salmon, Scaled H. E. tests, *Stanford Research Inst. Rept. 50*, November 1951.
- W. I., and T. C. Atchison, Rock breakage by explosives, *U. S. Bur. Mines Rept. Invest.* 1957.
- Sum, S. A., and P. L. Wegkamp, Photographic earth motion study, Scooter event, final report, *Edgerton, Germeshausen & Grier, Inc., Rept. L-510*, February 1961.
- G. W., G. H. Higgins, and C. E. Violet, Underground nuclear detonations, *J. Geophys. Res.* 64, 1457-1470, 1959.
- I., and A. R. Miller, The detonation of explosives, *Proc. Roy. Soc. London, A*, 194, 107, 1948.
- J. G., Crater measurements, Operation Scooter, *Engineering Research and Development Rept. WT-1105*, July 1958.
- Maenchen, G., and J. H. Nuckolls, The UNEC code, paper J in *Proc. Geophys. Lab.-Lawrence Radiation Lab. Cratering Symposium, Lawrence Radiation Lab. Rept. UCRL-6438*, 1961.
- Martin, D. L., and W. J. Hinze, Energy partition of underground explosions, *Engineering Research and Development Lab. Rept. AFSWP-789*, March 1958.
- Murphey, B. F., and L. J. Vortman, *J. Geophys. Research*, 66(10), 1961.
- Panama Canal Company, Crater tests in basalt, *Isthmian Canal Studies Memo 284-P*, Diablo Heights, Canal Zone, April 26, 1948a.
- Panama Canal Company, Crater tests in Gatun sandstone, *Isthmian Canal Studies Memo 285-P*, Diablo Heights, Canal Zone, May 14, 1948b.
- Panama Canal Company, Crater tests in marine muck, *Isthmian Canal Studies Memo 286-P*, Diablo Heights, Canal Zone, May 6, 1948c.
- Pokrovskii, G. I., and I. S. Fedorov, Effect of shock and explosion on deformable media, *Gos Izd.*, Moscow, USSR, 1957.
- Sachs, D. C., and L. M. Swift, Small explosion tests, project Mole, final report, *Stanford Research Inst. Rept. AFSWP-291*, December 1955.
- Shelton, A. V., M. D. Nurdyke, and R. H. Goeckermann, The Neptune event, *Lawrence Radiation Lab. Rept. UCRL-5766*, April 1960.
- Shoemaker, E. M., Impact mechanics at Meteor Crater, Arizona, *U. S. Geol. Surv. Open File Rept.*, July 1959.
- Thompson, T. L., and J. B. Misch, Geologic studies of underground nuclear explosions Rainier and Neptune, final report, *Lawrence Radiation Lab. Rept. UCRL-5757*, October, 1959.
- Vaile, R., *J. Geophys. Research*, 66(10), 1961.

(Manuscript received July 6, 1961.)

A Generalized Empirical Analysis of Cratering¹

CHARLES E. VIOLET

*Lawrence Radiation Laboratory, University of California
Livermore, California*

Abstract. A general empirical analysis of chemical and nuclear explosive cratering is presented. This analysis makes use of the concepts of efficiency and scaling formalism. Efficiency is assumed to depend on type of explosive, medium, and depth of burst. Scaling formalism is generalized in terms of experimentally determined yield exponents associated with each pertinent dimension. Experimental procedures are specified that determine the efficiency and yield exponents. Applying these methods to chemical explosive cratering data in desert alluvium, the crater radius and depth yield exponents are both $1/3.4$. The standard deviation is 3 per cent. The depth-of-burst yield exponent is $1/3.6$ with a standard deviation of 5 per cent. Thus the principle of similitude is not rigorously obeyed. Assuming that nuclear craters are best described by the above exponents, the percentage efficiency of nuclear cratering in desert alluvium is determined as follows:

	Based on Crater Radius	Based on Crater Depth
Jangle S	2.6 ± 1.2	4.6 ± 2.1
Jangle U	78 ± 35	146 ± 66
Teapot Ess	43 ± 19	208 ± 94

The behavior of the efficiency indicates that the relative contributions of various crater-forming mechanisms differ in chemical and nuclear cratering. Therefore the cratering capabilities of nuclear explosives cannot be related to those of chemical explosives by means of a single parameter. The prediction of nuclear crater dimensions from data obtained from low-yield chemical explosives is examined. The prediction error depends on yield extrapolation and depth-of-burst as well as the errors in the yield exponents and efficiency. For depths of burst near the surface or near the maxima of the depth-of-burst curves the prediction error is a minimum. The accuracy of the depth-of-burst yield exponent q is much less important than that of the crater dimension exponents p_i . The relationships between the precision that should be attempted for p_i and that for E is

$$\sigma_{p_i} = \sigma_E / |\ln EW|$$

where EW is the yield extrapolation. For yield extrapolations of 10^6 the accuracy of p_i should be an order of magnitude better than that of E .

1. INTRODUCTION

The largest amount of experimental information on crater dimensions from chemical explosives is from [Pokrovskiy, 1957; Murphey and Vortman, 1959; Murphey and McDougall, 1959; U. S. Army Corps of Engineers, 1958; Vortman and Vortman, 1959; and Vortman, Chabai, Perret, and Vortman, 1960]. Some nuclear cratering data in desert alluvium (John G. Lewis, private communication, 1958), wet coral sand (R. B. Vaile, private communication, 1955), and Oak Springs (H. J. Melton, Nordyke, and Goeckermann, 1960),

are also available. This information has generally been described and interrelated in terms of two concepts, scaling (principle of similitude) and efficiency.

Cube-root scaling [Chabai and Hankins, 1960], states that, for a limited range of energy release (yield) and for a given medium, crater dimensions can be described by the following expression:

$$R/W^{1/3} = f(H/W^{1/3}) \quad (1)$$

Presented at the Geophysical Laboratory-Lawrence Radiation Laboratory Cratering Symposium held at the Geophysical Laboratory in Washington, D. C., on March 28 and 29, 1961.

However, it is well known that cube-root scaling fails to predict crater dimensions accurately [Pokrovskiy, 1957]. Previous investigators have sought to improve on cube-root scaling by retaining similitude and deriving a so-called

'empirical scaling exponent.' We take a more general approach here in which similitude itself can be tested.

Although cube-root scaling is an inadequate description of cratering, the formalism of equation 1 is useful in an empirical analysis. Thus the exponents in equation 1 can be replaced by adjustable parameters which are determined experimentally. Since these parameters are not properly scaling exponents we refer to them as yield exponents.

To compare the cratering capabilities of nuclear and chemical explosives, the concept of efficiency has been used. (In this paper, 'efficiency' will refer to the cratering capabilities of nuclear explosives relative to chemical explosives, or a reference chemical explosive.) To attempt to relate chemical and nuclear cratering by means of this single parameter would appear to be unreasonable in view of the vast differences in the energy densities of these explosives. Also, for surface bursts, in which an appreciable fraction of the nuclear energy release is lost by radiation, the generation of the air and ground shock waves, as well as their configurations and interactions, differ greatly for the two explosives. For surface bursts, then, a comparison of cratering capabilities of chemical and nuclear explosives would probably require two parameters rather than one: one for crater radius and another for crater depth. If we make the comparison at depth rather than for surface explosions, however, a plausible comparison could perhaps be made using one parameter.

This might be possible in terms of a phenomenon known to occur in underground nuclear explosions: the formation of the 'gas ball.' Behind the outward-moving shock wave, a region filled with gas at extremely high temperature and pressure forms [Johnson, Higgins, and Violet, 1959]. This region exists for a characteristic time in a spherical volume of characteristic radius [Kennedy and Higgins, 1958]. In Oak Springs tuff this radius is approximately $10W^{1/3}$ feet [Johnson, Higgins, and Violet, 1959], where W is in kilotons. (A nuclear yield of 1 kiloton is defined as the prompt energy release of 10^{12} calories.) This radius would probably not differ markedly (less than a factor of 2) for other earth materials. A gas ball of characteristic radius also exists for chemical explosives, namely the radius of the explosive itself. In the same

units as above, this radius is approximately $15W^{1/3}$.

If the nuclear gas ball breaks above the surface, a larger fraction of its energy can be lost by radiative processes than in the equivalent chemical-explosive case, owing to the higher temperature of the initial nuclear gas ball. As the depth of burst increases, the fraction of the nuclear energy release available for crater production increases from its surface value and levels off. This plateau value is associated with a critical depth of burst which is approximately equal to the gas-ball radius. We refer to this plateau value as 'the efficiency.' Another process that affects the efficiency is the production of noncondensing gases (e.g., water vapor) associated with the high temperature and pressure of the nuclear gas ball. Thus, in this model, efficiency has a depth and medium dependence.

The mechanisms for craters produced by surface bursts are crushing, compaction, and other nonelastic processes. For greater depths of burst, which are of interest in this paper, the dominant mechanism is the expansion and bursting of the gas ball. A striking qualitative confirmation of this statement is obtained from the crater photography (Fig. 1, photographs by Edgerly, Germeshausen, and Grier, Inc.) of the Sedan event [Nordyke, 1961]. Although the yield exponents might have an anomalous behavior for detonations near the surface, at depths greater than the gas-ball radius the yield exponents would be expected to behave as 'slowly varying constants.' For the range of yields considered here we assume these exponents to be constant. (Evidence from Baldwin [1949] indicates this to be a good assumption over a yield range of at least three orders of magnitude.)

The gas ball for nuclear explosives is largely composed of vaporized earth materials. Since the vaporization of the materials involves irreversible processes, a smaller fraction of the total yield is available for crater production for nuclear explosives than for chemical explosives. Therefore, to a first approximation, the mechanism of nuclear cratering should be the same as that of a correspondingly smaller chemical explosion. For this model the efficiency is expressed as the coefficient of the prompt nuclear yield. We tentatively assume this to be a correct model. Its validity will be explored in this paper.

The current state of knowledge with respect

chemical and nuclear cratering can be summarized as follows:

the principle of similitude is retained, the value of the so-called 'empirical scaling factor' is $1/3.4$ for both chemical and nuclear explosives.

the cratering efficiency of tamped nuclear

explosives is somewhere between 20 and 150 per cent (John G. Lewis, private communication, 1958) [Chabai, 1959]. This range is apparently due to differences between chemical and nuclear cratering processes, in the emplacement and tamping conditions, and in the definitions of efficiency by previous investigators.

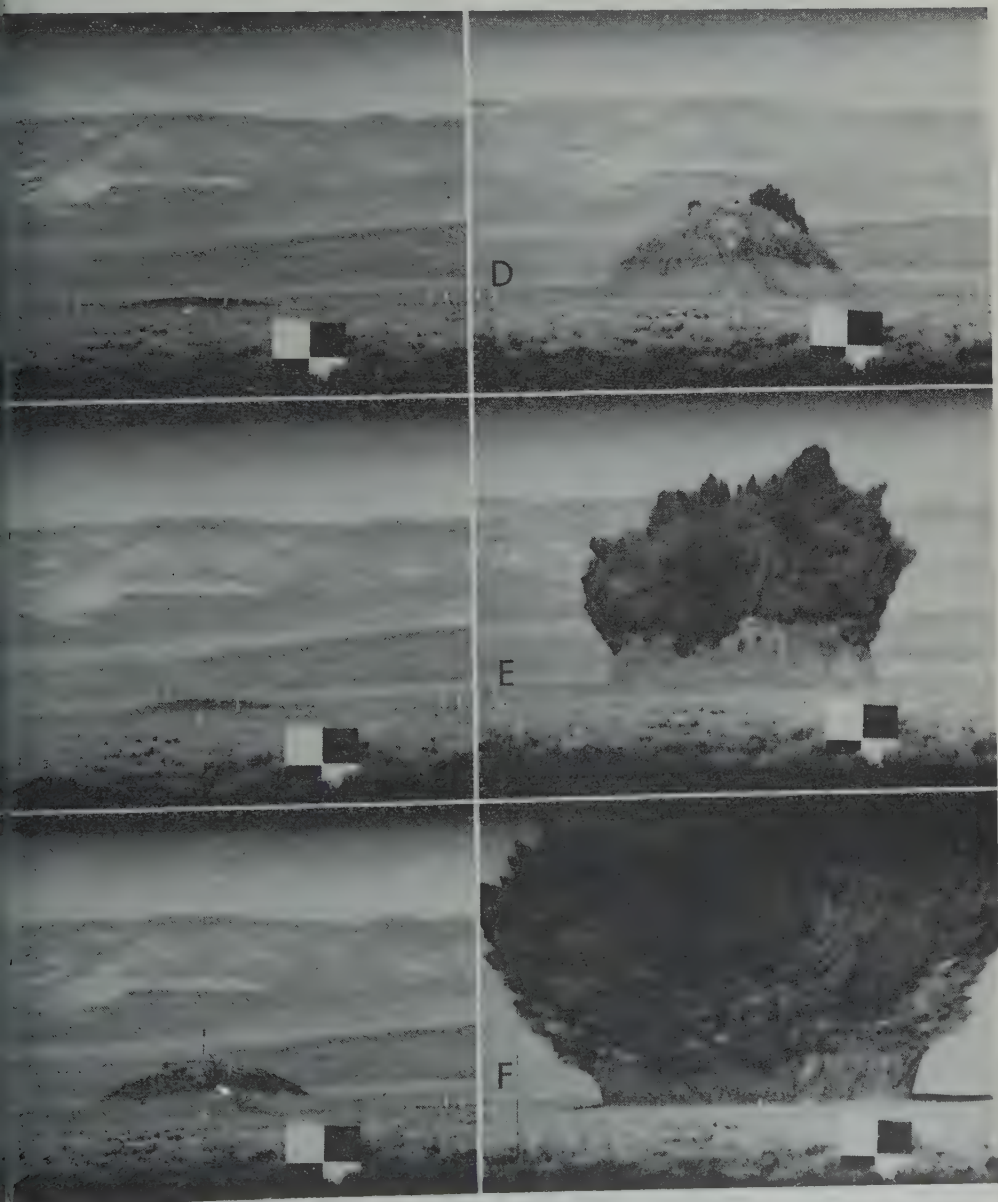


Fig. 1. A. Ground zero at 292 msec after detonation. B. Ground zero at 408 msec after detonation. C. Ground zero at 732 msec after detonation. D. Ground zero at 1092 msec after detonation. E. Ground zero at 1327 msec after detonation. F. Ground zero at 2233 msec after detonation.

3. The efficiency for crater depth is reported to be greater than that for crater radius by about a factor of 2 [Chabai, 1959].

From these considerations it can be concluded that the current status of cratering as a science or technology is in an unsatisfactory condition for the following reasons:

1. A theoretical treatment of cratering from first principles should be attempted. Some progress in this area is only recently apparent [Brode and Bjork, 1960].

2. All previous treatments of cratering have assumed the validity of the principle of similitude. However, since cube-root scaling is known to be inadequate the attendant failure of similitude would not be surprising. The analysis presented here provides a test of this principle by assuming an independent yield exponent associated with each dimension.

3. A meaningful definition of efficiency and a consistent method of its calculation should be adopted.

4. Experimental errors have received little attention either in experimental design or in the analysis of data. Consequently the uncertainty in predicting nuclear crater dimensions from low-yield chemical explosive cratering is not well known.

In an attempt to meet some of these problems, we shall discuss: (1) a general empirical analysis of cratering; (2) experimental procedures to provide for the determination of empirical parameters; (3) an application of this analysis to provide an empirical description of chemical and nuclear cratering in desert alluvium; (4) the effect of the errors in yield exponents and efficiency in predicting nuclear crater dimensions from low-yield chemical explosive cratering.

2. EMPIRICAL FORMALISM

We begin by stating our assumptions as discussed in the previous section.

1. Chemical and nuclear explosive cratering are related by means of one parameter, 'the efficiency.' It has low value for surface bursts, and with increasing depth of burst it increases and levels off to a plateau value. It is also a function of medium. It is mathematically expressed as a coefficient of the total nuclear yield.

2. A yield exponent is associated with each

dimension (crater radius, depth, and depth of burst). Yield exponents are independent of explosive (chemical or nuclear) and are constant for a given medium.

Generalizing equation 1 in accordance with these assumptions, we have

$$R_i^{c,n}/(EW)^{p_i} = f_i[H^{c,n}/(EW)^q]$$

where i denotes a specific crater dimension. For example, we let R_1 be crater radius and R_2 crater depth. p_i and q are the generalized yield exponents. The superscripts c and n refer to chemical or nuclear explosives. The curves (depth-of-burst curves) are determined by fitting convenient analytic functions to the experimental points.

A. Determination of yield exponents. Equation 1 can be expressed as

$$H^{c,n}/(EW)^q = F_i[R_i^{c,n}/(EW)^{p_i}]$$

To effect a determination of the scaling exponents, the observation equations 2 and 3 can be written

$$\begin{aligned}\ln R_i^{c,n} &= p_i \ln (EW) \\ &\quad + \ln f_i[H^{c,n}/(EW)^q] \\ \ln H^{c,n} &= q \ln (EW) \\ &\quad + \ln F_i[R_i^{c,n}/(EW)^{p_i}]\end{aligned}$$

If a set of observations could be obtained at constant values of the arguments of f_i and F_i , p_i and q could be readily obtained. In general it is impossible to choose data for constant arguments of f_i and F_i since the scaling exponents are unknown a priori. By a judicious choice of experimental conditions, this can be done, as is illustrated in the following section.

B. Determination of p_i . For the purpose of yield-exponent determination we define an 'effective surface' at an actual depth of $\approx 15W$ feet. (The best a priori value of q should be chosen. On the basis of previous work we choose $q = 1/3.4$. From equation 14, Figure 7, and a discussion of errors in section 4 it is evident that for depths of burst near the effective surface the systematic error arising from this choice of q is much less than the random errors and can be neglected.) If we measure actual depths from this effective surface, equation 4 becomes

$$\ln R_i^c = p_i \ln W + \ln A_{i0}$$

$$\ln R_i^n = p_i \ln W + (p_i \ln E + \ln A_{i0})$$

A_{i0} are constants (the first terms in series expansions of f_i). Thus a plot against $\ln W$, as obtained from chemical explosions at the 'effective surface,' be consistent with straight lines through $\ln h$. The values of p_i and E can be obtained from the slopes and intercepts of these curves.

Determination of depth-of-burst scaling exponents. Knowledge of the crater-dimension exponents permits the subsequent determination of the depth-of-burst scaling exponent. Consider first two series of chemical experiments at various depths, one at a yield of W_1 and the other at W_2 . (conditions

$$R_{i,1}^c/R_{i,2}^c = (W_1/W_2)^{p_i} \quad (7)$$

from equation 3 that

$$H_1^c/H_2^c = (W_1/W_2)^q \quad (8)$$

Knowledge of the depth-of-burst ratio for scaled crater dimensions permits solving for the depth-of-burst scaling exponent, q , from equation 5.

A similar analysis could be carried out for explosive cratering if appropriate data are available. Equations 7 and 8 are valid for explosive as well as for chemical explosions if the yield is constant. Thus, to determine q for explosive cratering, depths of burst deeper than the critical depth must be used.

D. Determination of efficiency for general experimental conditions. Although E can be determined under the limited conditions of paragraph B a more general method is necessary. Knowledge of the yield exponents permits the general determination of the efficiency if the depth-of-burst curves are expressed analytically. The functions f_i can be expressed as power series of the scaled depth of burst, h , where

$$h = [H^{c,n}/(EW)^a]$$

Thus:

$$f_i = \sum_{m=0}^{m=M} A_{im} h^m \quad (9)$$

Equation 2 then becomes

$$R_{i,1}^{c,n} = (EW)^{p_i} \cdot \sum_{m=0}^{m=M} A_{im} h^m \quad (10)$$

Carrying out the indicated multiplication, we have

$$R_{i,1}^c = A_{i0} W^{p_i} + A_{i1} H W^{p_i-q} + A_{i2} H^2 W^{p_i-2q} + \dots \quad (11a)$$

and

$$R_{i,1}^n = A_{i0} (EW)^{p_i} + A_{i1} H (EW)^{p_i-q} + A_{i2} H^2 (EW)^{p_i-2q} + \dots \quad (11b)$$

We note that, if second-order and higher

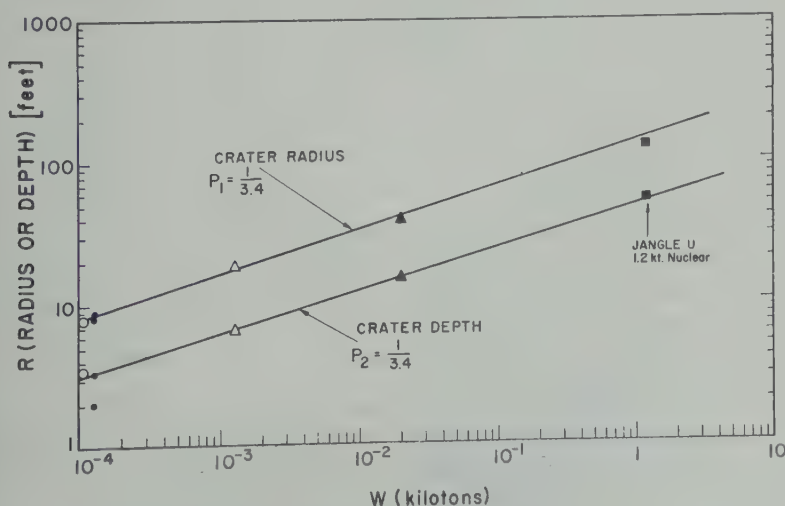


Fig. 2. Crater dimensions for chemical explosions of scaled depth of burst from 12 to 18 ft/kt^{1/3.4} in desert alluvium.

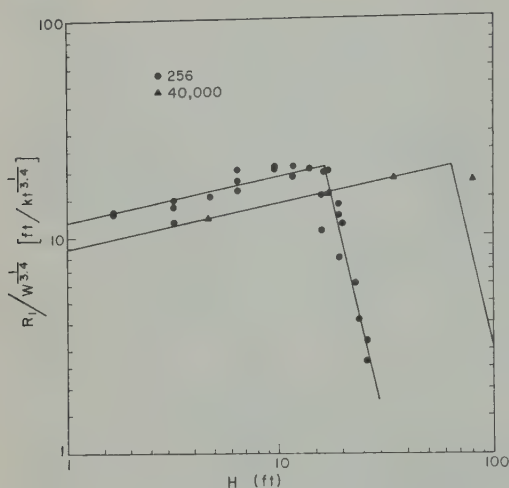


Fig. 3. Scaled crater radii for chemical explosions plotted as a function of absolute depth of burial.

terms are dropped, and $p_i \approx u$, equations 11 reduce to

$$R_i^c = A_{i0}W^{p_i} + A_{i1}H \quad (12a)$$

$$R_i^n = A_{i0}(EW)^{p_i} + A_{i1}H \quad (12b)$$

For equal crater dimensions and equal depths of burst, the efficiency in this approximation is the ratio of chemical to nuclear yields.

This statement has been given as a definition of efficiency [Chabai, 1959; Vortman, 1959]. It is, at best, however, only a first-order approximation. Knowledge of scaling exponents and the coefficients A_{im} in a given medium as obtained from chemical-cratering data permits solving equation 11b for E from any set of nuclear-cratering data in that medium.

3. AN ANALYSIS OF CHEMICAL AND NUCLEAR CRATERING IN DESERT ALLUVIUM

The only material for which direct nuclear scaling information is available is saturated coral sand (R. B. Vaile, private communication, 1955). Unfortunately, these data are of questionable value because (1) there have been no detonations below the surface of the earth; (2) sometimes there were several feet of water between the center of detonation and the coral sand; (3) the craters were eroded by sea water before measurement; (4) the properties of the material changed to some extent with depth. The shallow craters from low-yield detonations were formed in coral

sand. The deep craters from high yields occurred in partly cemented coral rock.

The nuclear crater in Oak Springs tuff is a limited value since it occurred on a hillside with a slope of about 30° [Shelton, Nordyke, and Goeckermann, 1960].

Desert alluvium is the only material for which unambiguous nuclear cratering data are available. Unfortunately, all nuclear craters in the medium were produced from detonations of the same yield. Thus assumption 1 of section 2 can be tested directly, but assumption 2 cannot be tested for nuclear cratering. There is also an abundance of chemical explosive crater data in desert alluvium obtained from center-detonated spherical charges. For these reasons we shall analyze cratering data in desert alluvium in order to exemplify the methods of the previous section. A tabulation of chemical and nuclear cratering data in desert alluvium is given by Nordyke [1961].

A. Determination of p_i . Crater dimensions for chemical explosions of scaled depths in the range of 12–18 ft/kt $^{1/3.4}$ are plotted in Figure 3. The best fit for the chemical crater dimensions corresponds to $p_1 = p_2 = 1/3.4$. The associated standard deviation is 3 per cent. Crater dimensions for Jangle U, where the scaled depths are 16 ft/kt $^{1/3.4}$, are also plotted in Figure 2. The nuclear data appear to be consistent with the curves determined from the chemical data.

B. Determination of q . $R_1/W^{1/3.4}$ and $R_2/W^{1/3.4}$ for 256 and 40,000-pound data are plotted a

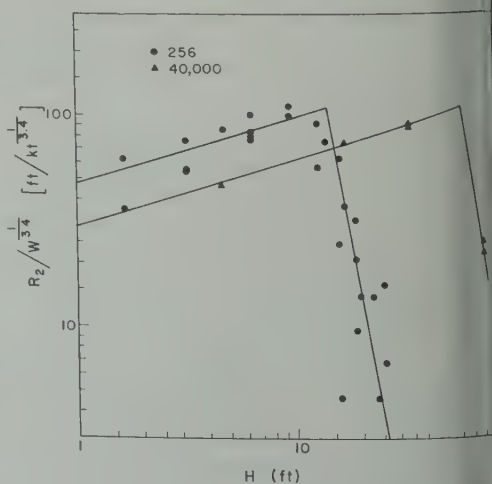


Fig. 4. Scaled crater depth for chemical explosions plotted as a function of absolute depth of burial.

of absolute depth (H) in Figures 3 and 4. The data of the form of equation 9 could be fitted to the data; for the sake of clarity and convenience, however, we fit curves to these data of the type suggested by Chabai [1959]. Using equations 4 and 5, we find that for both plots of $R_1/W^{1/3.4}$ and $R_2/W^{1/3.4}$ as a function of $H/W^{1/3.6}$ in Figures 5 and 6, thus obtaining the depth-of-burst f_i . The coefficients for the first four terms

in the expansion of f_i are given in Table 1. The errors are standard deviations.

Using the experimentally determined exponents, the coefficients of Table 1, and the crater dimensions for Jangle S, Jangle U, and Teapot Ess [Nordyke, 1961], we calculate the efficiencies shown in Table 2. The errors are standard deviations (see section 4).

Equation 12b is sufficient for Jangle S and Jangle U, but second-order terms (equation 11b) must be included for Teapot Ess.

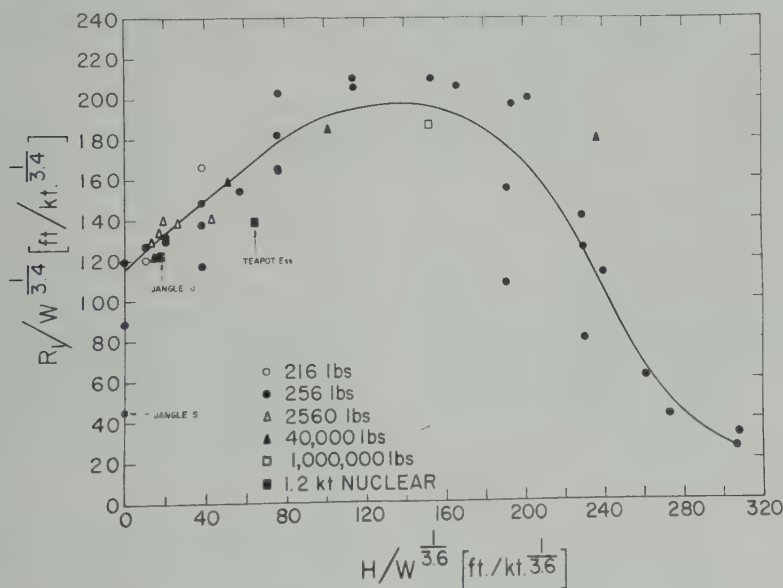


Fig. 5. Scaled crater radii as a function of scaled depth of burst for desert alluvium.

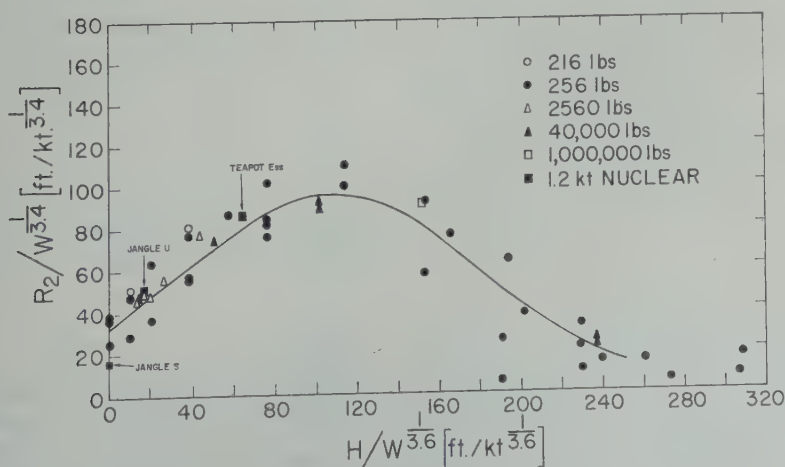


Fig. 6. Scaled crater depth as a function of scaled depth of burst for desert alluvium.

TABLE 1. Coefficients for Desert Alluvium

$A_{10} = 116 \pm 9$	$A_{20} = 33 \pm 4$
$A_{11} = 0.9 \pm 0.1$	$A_{21} = 0.8 \pm 0.1$
$A_{12} = -(1.5 \pm 0.5)$	$A_{22} = -(1.7 \pm 0.6)$
$\times 10^{-3}$	$\times 10^{-3}$
$A_{13} = -(6 \pm 3) \times 10^{-6}$	$A_{23} = -(7 \pm 3) \times 10^{-6}$

A statistical test of the values of Table 2 by means of Student's *t*-distribution [Hoel, 1954] shows that the differences in the efficiencies of Table 2 are significant with the possible exception of those based on crater depth for Jangle U and Teapot Ess. We take this to be a real effect in accordance with the conclusions of previous investigators (John G. Lewis, private communication, 1958), [Chabai, 1959]. Whereas this behavior of efficiency based on crater radius is consistent with the simple model of section 1, that based on crater depth is not. Furthermore, its dependence on crater dimension is not consistent with this model. Thus, the relative contributions of various crater-forming mechanisms, e.g., gas-ball expansion, nonelastic deformation, and subsidence, differ in the chemical and nuclear cases, and the cratering capabilities of nuclear and chemical explosives cannot be related by a single parameter.

4. UNCERTAINTIES IN CRATER-DIMENSION PREDICTIONS

We shall examine the uncertainties in crater-dimension predictions for nuclear craters in the 10⁶- to 10⁸-ton yield range based on chemical-cratering data obtained in the 10⁻¹ to 10-ton range. Thus the typical yield extrapolation is a factor of 10⁶. A discussion of the relative contribution of the errors in efficiency and scaling exponents to the net error follows.

The net standard deviation associated with a predicted crater dimension is

$$\Sigma_{R_i}^2 = \sigma_{R_i}^2 + \sigma_e^2$$

where

Σ_{R_i} = the net standard deviation for a predicted crater dimension.

σ_{R_i} = the standard deviation in R_i resulting from the errors in efficiency and yield exponents (prediction error).

σ_e = the standard deviation of a single crater observation.

Since σ_e can be obtained from chemical explosion of a given yield, σ_e and σ_{R_i} are uncorrelated; we restrict the discussion to an evaluation of σ_{R_i} for desert alluvium. From equation 2 we derive

$$\begin{aligned} dR_i^{c,n}/R_i^{c,n} = & (p_i + E/f_i \partial f_i/\partial E) dE/E \\ & + (p_i \ln EW) dp_i/p_i \\ & + (q/f_i \partial f_i/\partial q) dq/q \end{aligned}$$

In terms of standard deviations we have

$$\begin{aligned} \sigma_{R_i}^2 = & (p_i + \epsilon_i(h)q)^2 \sigma_E^2 + (p_i \ln EW)^2 \sigma_{p_i}^2 \\ & + (\epsilon_i(h)q \ln EW)^2 \sigma_q^2 \end{aligned}$$

where σ_{R_i} , σ_E , σ_{p_i} , and σ_q are the fractional standard deviations in $R_i^{c,n}$, E , p_i , and q respectively, and

$$\begin{aligned} \epsilon_i(h) = & 1/q E/f_i \partial f_i/\partial E \\ = & 1/\ln EW \ 1/f_i \partial f_i/\partial E \end{aligned}$$

Evidently the prediction error depends on yield extrapolation and depth of burst as well as errors in p_i , q , and E .

The dimensionless quantity $\epsilon_i(h)$ is plotted in Figure 7 in the case in which $f_i(h)$ is represented by equation 9 and the coefficients $A_{i,m}$ are taken from Table 1.

For depths of burst near the surface or near the maxima of the depth-of-burst curves the following condition holds:

$$|\epsilon_i(h)| \ll 1$$

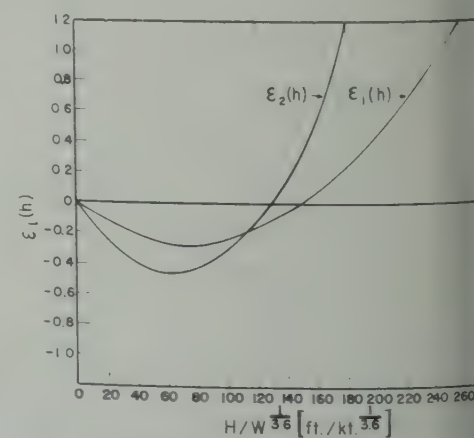


Fig. 7. The function ϵ_i (see text) plotted versus scaled depths of burst.

TABLE 2. Per Cent Efficiencies for Desert Alluvium

	Based on Crater Radius	Based on Crater Depth
S	2.6 ± 1.2	4.6 ± 2.1
U	78 ± 35	146 ± 66
Ess	43 ± 19	208 ± 94

Figs. 5, 6, and 7).

is condition and with the aid of equation observe that E and q are least accurately ded, and that the prediction error is a

n. ext calculate the error in E . From equation and under the condition of equation 15 is valid for the nuclear data in desert we derive

$$1/p_i^2 [\sigma_e^2 + (p_i \ln EW)^2 \sigma_{p_i}^2] \quad (16)$$

ves of Figures 5 and 6 are essentially data from 256-pound explosions. Thus appropriate yield extrapolation is $\approx 10^4$. the values for p_i and σ_{p_i} from section 3 = 0.10 (see the spread of data in Figs. we have

$$\sigma_E \approx 45 \text{ per cent}$$

these values we can now calculate the on error for a yield extrapolation of 10^4 the condition of equation 15 from 14. Thus $\sigma_{Ri} \approx 17$ per cent. Therefore standard deviation for a predicted nuclear dimension in desert alluvium involving a extrapolation of 10^4 is

$$\Sigma_{Ri} \approx 20 \text{ per cent}$$

n experiment designed to provide data dicting nuclear crater dimensions in a aterial the following relationship among parameter uncertainties should hold under dition of equation 15:

$$|\sigma_{p_i} \ln EW| \approx |\sigma_e \epsilon_i(h) \ln EW| \quad (17)$$

raw the following conclusions from equa- (1) the accuracy of the depth-of-burst onents (q) is much less important than the crater-dimension yield exponents; (2) onship between the precision that should

be attempted for p_i and the precision of E so that approximately equal uncertainties will be introduced in the calculations of R_i is given by

$$\sigma_{p_i} \approx \sigma E / |\ln EW| \quad (18)$$

Clearly, the precision of the yield exponents is much more important than that of efficiency. For our assumed extrapolation range of 10^4 in yield, the accuracy of p_i should be more than an order of magnitude better than that of E . In desert alluvium the scaling exponents are known to an accuracy of about 3 per cent. To explore cratering phenomena in another medium to this same extent would require a knowledge of the efficiency to about a factor of $1/3$. Thus, it would appear that medium effects could be adequately explored using chemical explosives only.

5. CONCLUSIONS

A. A general empirical analysis such as that presented here provides a powerful tool for describing and interrelating chemical and nuclear explosive cratering. Such concepts as the principle of similitude and the efficiency as represented by one parameter can be readily tested by this analysis.

B. The analysis of chemical cratering in desert alluvium gives the following yield exponents for yields in the range 256 to 40,000 pounds.

$$1/p_1 = 3.4 \text{ (crater radius)}$$

$$1/p_2 = 3.4 \text{ (crater depth)}$$

$$1/q = 3.6 \text{ (depth of burst)}$$

The standard deviation associated with p_1 and p_2 is 3 per cent. That for q is 3 per cent. For these phenomena, then, the principle of similitude is not rigorously obeyed.

C. Assuming that nuclear cratering is best described by the same yield exponent as chemical cratering, the per cent efficiencies for nuclear cratering are

	Based on Crater Radius	Based on Crater Depth
Jangle S	2.6 ± 1.2	4.6 ± 2.1
Jangle U	78 ± 35	146 ± 66
Teapot Ess	43 ± 19	208 ± 94

Thus the relationship between chemical any nuclear cratering capabilities (the efficiency) cannot be given in terms of one parameter.

D. The prediction error depends on yield extrapolation and depth of burst as well as the errors in the yield exponents and efficiency. For depths of burst near the surface or the maxima of the depth of burst curves: (1) E and q are least accurately determined; (2) the error in predicting nuclear crater dimensions from low-yield chemical explosives is a minimum; (3) the net standard deviation for a predicted nuclear crater dimension in desert alluvium involving a yield extrapolation of 10^6 is approximately 20 per cent.

E. The relationship between the uncertainties in cratering parameters is summarized as follows: (1) The accuracy of the depth-of-burst yield exponent q is much less important than that of the crater dimension yield exponents p_i . (2) The relationship between the precision that should be attempted for p_i and the precision of E is given by

$$\sigma_{p_i} = \sigma_E / |\ln EW|$$

where EW is the yield extrapolation.

F. For yield extrapolations of the order of 10^6 or more the accuracy of p_i should be more than an order of magnitude better than that of E . For these conditions, medium effects could be adequately explored with chemical explosives.

Acknowledgments. I am indebted to Drs. G. W. Johnson and M. D. Nurdyke for their helpful advice and encouragement. The preparation of the manuscript by Mrs. Wilma McGurn is gratefully acknowledged.

This work was performed under the auspices of the U. S. Atomic Energy Commission.

REFERENCES

- Baldwin, R. B., *Face of the Moon*, University of Chicago Press, Chicago, 1949.
- Brode, H. L., and R. L. Bjork, Cratering from megaton surface burst, *Rand Corp. RM* 1240, June 30, 1960.
- Chabai, A. J., Crater scaling laws for desert alluvium, *Sandia Corp. SC 4391 (RR)*, December 1960.
- Chabai, A. J., and D. M. Hankins, Gravity scaling laws for explosion craters, *Sandia Corp. SC 4486 (RR)*, December 1960.
- Hoel, Paul G., *Introduction to Mathematical Statistics*, 2nd ed., John Wiley & Sons, New York, 1954.
- Johnson, G. W., G. H. Higgins, and C. E. Vortman, Underground nuclear detonations, *J. Geophysical Research*, 64(10), 1457-1470, 1959.
- Kennedy, G. C., and G. H. Higgins, Temperatures and pressures associated with the cavity produced by the Rainier event, *UCRL-5281*, July 1959.
- Murphey, B. F., and H. R. MacDougall, Cratering studies in desert alluvium, *Sandia Corp. SC 4119-59 (51)*, May 1959.
- Murphey, B. F., and L. J. Vortman, Cratering with chemical explosives, *Proc. Second Plowshare Symposium*, May 13-15, *UCRL-5676*, 5-19, 1959.
- Nurdyke, M. D., On Cratering, *UCRL 6578*, August 1961.
- Pokrovskiy, G. I., Diffusion blasting: Coal mining, *Military Aeronautical Engineering Academy, Moscow, Prioroda*, no. 8, 81-83, August 1958.
- Shelton, A. Vay, M. D. Nurdyke, and R. M. Goeckermann, The Neptune event, A nuclear cratering experiment, *UCRL-5766*, April 1959.
- U. S. Army Corps of Engineers, Cratering effects of surface and burial H. E. charges in loess and clay, *Tech. Rept. 2-482*, Waterways Experiment Station, Vicksburg, Mississippi, June 1958.
- Vortman, L. J., Relative cratering efficiency of nuclear explosives, *Sandia Corp. SCTM 1147 (51)*, Albuquerque, New Mexico, April 1959.
- Vortman, L. J., and L. N. Schofield, High-explosive cratering in Fan delta alluvium, *Sandia Corp. SCTM 60-59 (51)*, October 1959.
- Vortman, L. J., A. J. Chabai, W. R. Perret, and J. W. Reed, Project Buckboard, Interim Report, *Sandia Corp. SC 4486 (RR)*, November 1959.

(Manuscript received July 6, 1961.)

Experimental Determination of Earthquake Fault Length and Rupture Velocity¹

FRANK PRESS, ARI BEN-MENAHEN, AND M. NAFT TOKSÖZ

*Seismological Laboratory, California Institute of Technology
Pasadena, California*

Abstract. Three methods of determining the fault parameters of length and rupture velocity are combined with ultrasonic models. The theory behind the methods is shown to have a valid high approximate basis. Oversimplified assumptions and imperfect experimental data restrict the results to only rough indications of fault parameters. When applied to the great Chilean earthquake of May 1960, a fault length of the order of 1000 km and a rupture velocity near the speed of shear waves in crustal rock are found.

Introduction. A major problem of geophysics is to determine the nature of the earthquake mechanism. Although some theoretical speculations have been advanced, real progress must await more detailed descriptions of the strain fields and the rupture itself. Much has been learned by use of the radiation method to obtain fault plane parameters [Hodgson, 1957]. For example, it has been found that fault movements are predominantly of the strike-slip variety. Knopoff and Gilbert [1960] studied the theoretical basis of the radiation method on the assumption of radiation from a propagating finite fault. Recently, attempts have been made to equalize surface motions to provide additional information for fault solutions [Aki, 1960; Brune, Nafe, and Slichter, 1960].

In this paper we describe experimental methods and give some preliminary results of determining two additional parameters of the earthquake focus, namely fault length and rupture velocity. The methods are based on an analysis of seismograms of free oscillations of the earth and of long-period mantle Rayleigh and G waves. The theoretical bases for these methods have been described elsewhere [Press, and Smith, 1961; Ben-Menahem, 1961]. We shall review the theory, describe the model experiments which test the methods, and, finally, give results for the great Chilean earthquake of 1960. We make the basic

and reasonable assumption that the faulting process may be represented by a moving source. Not so reasonable is the assumption that the fault progresses only in one direction from the epicenter with constant velocity and strength. Since these last conditions are violated in actual earthquakes, the method will serve initially for giving rough indications of the fault parameters.

Fault parameters from free oscillations. The relative amplitudes and phases of the normal mode oscillations of the earth carry the necessary information for specifying the source, provided the earth's structure is known. In view of the large number of modes and overtones that occur and of the uncertainties that accompany amplitudes and phase determinations, this approach is a difficult one and has not yet been attempted. If we restrict ourselves to a determination of fault length and rupture velocity, a simple and more feasible method is available involving the phase shift between vertical and horizontal components of motion of each mode [Benioff, Press, and Smith, 1961].

For spheroidal free oscillations the phase shift is 0° or 180° . These oscillations may be thought of as the standing wave pattern formed by progressive Rayleigh waves of equal amplitude traveling in opposite directions. Rayleigh waves propagating in one direction are characterized by phase shifts of 90° . It is possible for a faulting source to excite free oscillations or progressive Rayleigh waves, depending on the relative amplitudes of the waves leaving the source in opposite directions. Hence the phase shift

¹Contribution 1040, Division of Geological Sciences, California Institute of Technology, Pasadena, California.

observed at a distant station is a measure of the source radiation pattern, which in turn depends on fault length and rupture velocity. This qualitative physical description of the method becomes increasingly valid for the higher modes ($n \geq 10$) where the free oscillations may be decomposed into traveling waves.

Ultrasonic model study. To illustrate the method and test its feasibility, we have made a two-dimensional model study using a circular plate and the equivalent of a moving source.

The steady-state expressions for free oscillations of a circular cylinder are

$$\begin{aligned} U_r &= -R(r) \cos n\theta \cos \omega t \\ U_\theta &= \Theta(r) \sin n\theta \cos \omega t \end{aligned} \quad (1)$$

where U_r and U_θ are displacements in the radial and polar directions r and θ , respectively, n is an integer which defines the mode number, and ω is circular frequency. The functions R and Θ which satisfy the equations of motion are

$$\begin{aligned} R &= \frac{A}{h^2} \frac{dJ_n(hr)}{dr} + \frac{Bn}{k^2} \frac{J_n(kr)}{r} \\ \Theta &= \frac{An}{h^2} \frac{J_n(hr)}{r} + \frac{B}{k^2} \frac{dJ_n(kr)}{dr} \end{aligned} \quad (2)$$

where $h = \omega/\alpha$, $k = \omega/\beta$, and α and β are the compressional and shear velocities, respectively. In the case of a cylindrical plate α is understood to be the plate velocity. A and B are amplitude factors whose ratio can be determined from either of the two boundary conditions of vanishing radial and tangential stress at the rim of the cylinder $r = a$. A and B may be eliminated between the two boundary equations to yield the frequency equation

$$\begin{aligned} &\left[(k^2 - 2h^2)J_n(ha) - 2 \frac{d^2 J_n(ha)}{da^2} \right] \\ &\cdot \left[\left(k^2 - \frac{2n^2}{a^2} \right) J_n(ka) + \frac{2}{a} \frac{dJ_n(ka)}{da} \right] \\ &- 4 \frac{n^2}{a^2} \left[\frac{dJ_n(ha)}{da} - \frac{J_n(ha)}{a} \right] \\ &\cdot \left[\frac{dJ_n(ka)}{da} - \frac{J_n(ka)}{a} \right] = 0 \end{aligned} \quad (3)$$

This equation is the same as the one for Rayleigh waves propagating along the circumference of a cylinder with phase velocity c [Ewing, Jardetzky,

and Press, 1957, p. 264] when $n = \omega a/c$. In this case, n assumes integral values only. Solution of (3) for ω vs. n can be obtained from the phase velocity curves computed by Oliver [1955].

We can represent a moving source by a succession of impulses shifted in space and delayed in time. When the impulse moves over an angular distance ϵ_0 with velocity c_0 , the displacement at a point θ on the rim is the superposition of displacements due to source positions between $\epsilon = 0$ and $\epsilon = \epsilon_0$ with proper delay $a\epsilon/c_0$:

$$\begin{aligned} U_r^{(n)} &= -R_n(a) \int_0^{\epsilon_0} \cos n(\theta - \epsilon) \\ &\quad \cdot \cos \omega_n \left(t - \frac{a\epsilon}{c_0} \right) d\epsilon \\ U_\theta^{(n)} &= \Theta_n(a) \int_0^{\epsilon_0} \sin n(\theta - \epsilon) \\ &\quad \cdot \cos \omega_n \left(t - \frac{a\epsilon}{c_0} \right) d\epsilon \end{aligned}$$

Performing these integrations we can write the displacements as

$$\begin{aligned} U_r^{(n)} &= -R_n(a) (P_n^2 + Q_n^2 \\ &\quad + 2P_n Q_n \cos \alpha_n)^{1/2} \cos (\omega_n t + \gamma_n + \varphi_r^{(n)}) \\ U_\theta^{(n)} &= -\Theta_n(a) (P_n^2 + Q_n^2 \\ &\quad + 2P_n Q_n \cos \alpha_n)^{1/2} \\ &\quad \cdot \cos (\omega_n t + \gamma_n - \varphi_\theta^{(n)} - \pi/2) \end{aligned}$$

where

$$\begin{aligned} P_n &= \frac{\sin [(\epsilon_0/2)(n - a\omega_n/c_0)]}{n - a\omega_n/c_0} \\ Q_n &= \frac{\sin [(\epsilon_0/2)(n + a\omega_n/c_0)]}{n + a\omega_n/c_0} \\ \alpha_n &= 2n(\theta - \epsilon_0/2) \\ \gamma_n &= (\epsilon_0/2)(n - a\omega_n/c_0) - n\theta \\ \varphi_r^{(n)} &= \tan^{-1} \left\{ \frac{Q_n \sin \alpha_n}{P_n + Q_n \cos \alpha_n} \right\} \\ \varphi_\theta^{(n)} &= \tan^{-1} \left\{ \frac{Q_n \sin \alpha_n}{P_n - Q_n \cos \alpha_n} \right\} \end{aligned}$$

The phase difference between displacements $U_r^{(n)}$ and $U_\theta^{(n)}$ for the n th mode is

$$\delta_n = \pi/2 + \varphi_r^{(n)} + \varphi_\theta^{(n)}$$

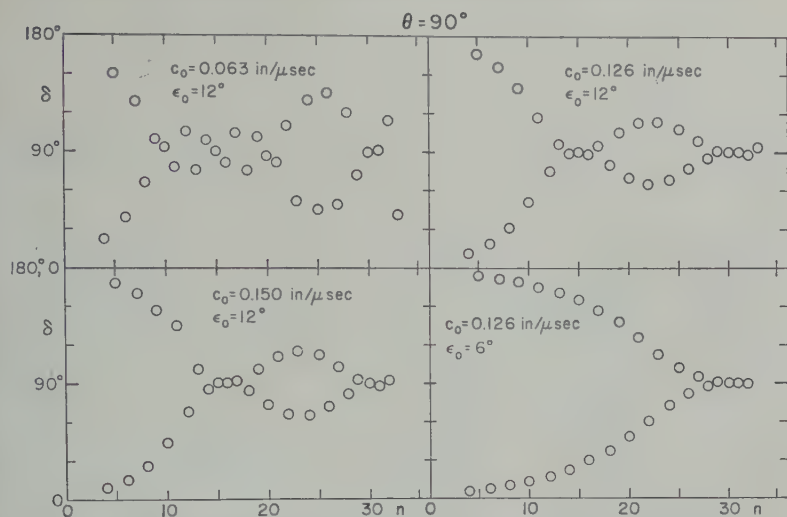


Fig. 1. Phase difference between vertical and horizontal displacement of free oscillations of cylindrical plate at distance 90° from source for various fault lengths ϵ_0 and rupture velocities c_0 .

tions of δ_n for values $n = 4$ to $n = 30$ and distances $\theta = 90^\circ, 120^\circ$, and 180° are given in Figs. 1, 2, 3, for several combinations of fault length ϵ_0 and rupture velocity c_0 . The results show interesting features. For n small, δ_n approaches the theoretical value of 0° which characterizes free oscillations of a plate by a nonmoving source or a point source. As n becomes large, δ_n is to be expected, for n small corresponds to wavelengths large compared with fault length, therefore the fault is effectively a point

source. Similarly, for small or intermediate values of n , δ_n approaches 0 or π as ϵ_0 becomes small or c_0 becomes large. For large n , δ_n hovers around 90° . This corresponds to a preferred radiation in a single direction from the source. In this case, progressive Rayleigh wave propagation predominates with the accompanying 90° phase shift. Free oscillations with 0° and 180° phase shifts correspond to the standing wave interference patterns between equal and opposite traveling Rayleigh waves.

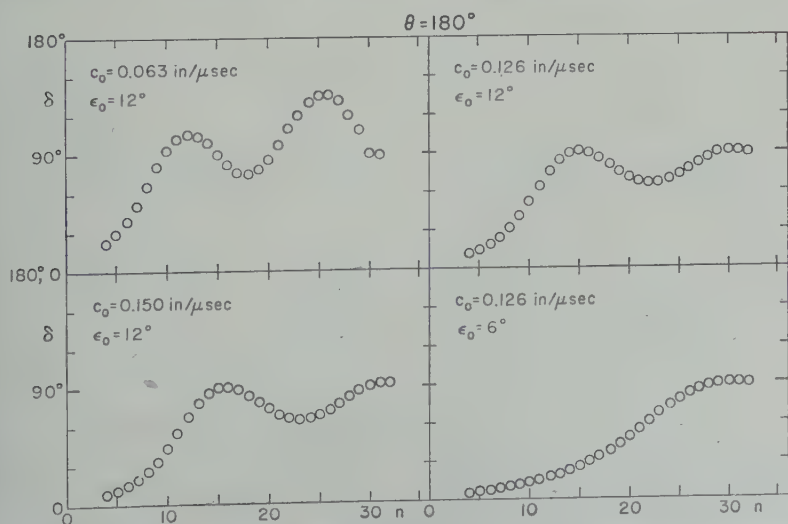


Fig. 2. Phase difference between vertical and horizontal displacement of free oscillations of cylindrical plate at distance 120° from source for various fault lengths ϵ_0 and rupture velocities c_0 .

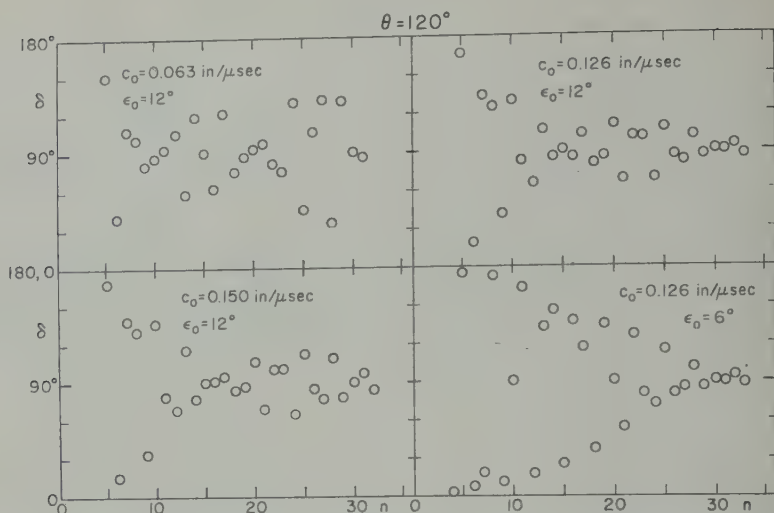


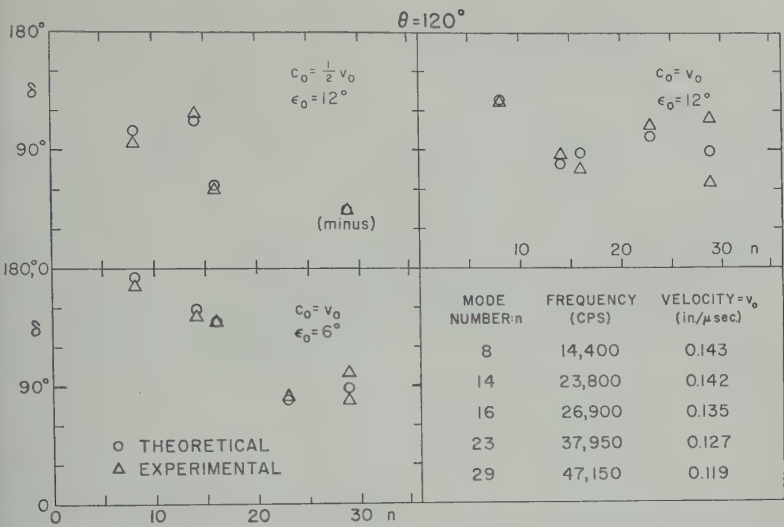
Fig. 3. Phase difference between vertical and horizontal displacement of free oscillations of cylindrical plate at distance 180° from source for various fault lengths ϵ_0 and rupture velocities c_0 .

In the ultrasonic model experiment a circular aluminum plate of 12-inch radius and 1/16-inch thickness was used. The experimental setup is essentially that described by *Healy and Press* [1960]. The source was a solid barium titanate cylinder, $1/8$ inch long and $1/8$ inch in diameter. The receiver was a barium titanate bimorph transducer.

The free modes of oscillation of the model plate were excited with steady-state sinusoidal source. This enabled us to work directly in frequency space without the need for Fourier analysis. The frequency of the source was varied until the natural frequency of the plate was found and resonance occurred. The nodes and antinodes were easily recognized along the rim, and mode number n was determined by counting the number of nodal points. The natural frequencies thus measured were compared with values calculated from Oliver's phase velocity curve and were found to agree to better than 1 per cent for modes lower than the 16th and better than 2 per cent for modes below the 30th. The frequencies derived from the free oscillations were systematically lower, a discrepancy which can be explained by Oliver's use of a Poisson's ratio of 0.25 instead of the 0.26 value of our model.

The moving source condition was established approximately. In the first place, the receiver was moved instead of the source. Such an interchange is very convenient for the experiment,

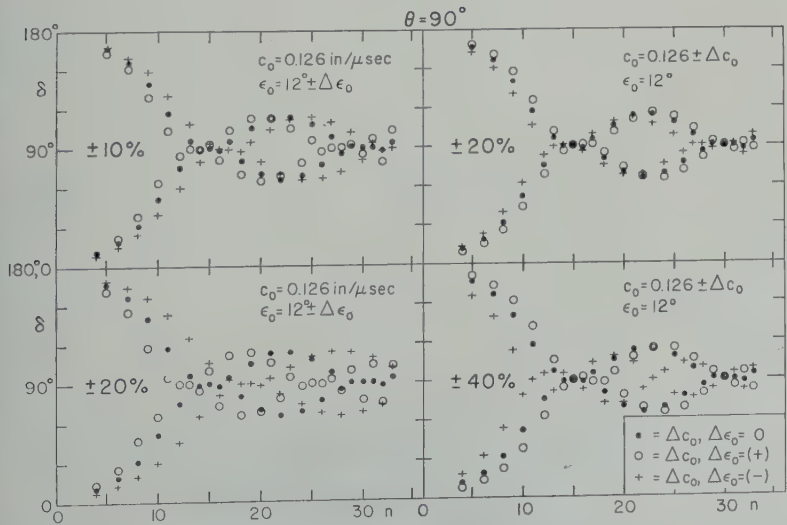
and is justified by the reciprocity theorem. The motion of the source was approximated by a series of discrete point sources spaced at 2° intervals and delayed in time. In our experiment, the readings were taken at 2° intervals by keeping the source fixed and shifting the receiver position 2° after each reading. The records obtained in this way were time shifted by an amount equal to the translation time of the source over a distance of 2° . The records were then superimposed to obtain the record equivalent to a moving source. Such a method is very practical because of its flexibility. However, it does not provide the precision we would like to have. It lacks the ability to excite an entire suite of free oscillations. A better method would be to place a desired number of source crystals at a succession of positions and to pulse them in successive order with proper delays. This technique requires a multichannel delaying system with fractional microsecond precision. We are building such an apparatus now. Seismograms of vertical and horizontal components of motion, recorded at 2° intervals, were digitized, time shifted, and added up. The phase difference between the resultant sinusoidal traces of vertical and horizontal components was measured. These measurements were made for five modes for two different source velocities and for two different source lengths for each mode. The velocities were chosen to be approximately equal



4. Comparison of theoretical and experimental values of phase differences for several modes n at distances $\theta = 120^\circ$ from source.

average and one-half of the average of the velocity C_R . The lengths were taken to ϵ_0 and $\epsilon_0 = 12^\circ$. The experimental values of differences are given in Figure 4, with theoretical values computed from 6. of the approximate method we have simulate a moving source, the agreement experiment and theory is reasonably however, if we wish to use this method

to determine the properties of the source, more precise experimental data are needed as well as data for a larger number of modes. This can be seen from Figure 5, in which we theoretically examine the sensitivity of the method. The effect of perturbations in fault length and rupture velocity is such that phase data for about 30 successive modes are required with a precision of $\pm 5^\circ$ if these fault parameters are to be determined with an accuracy of ± 10 to



5. The effect of perturbations in fault length ϵ_0 and rupture velocity c_0 on phase shift δ .

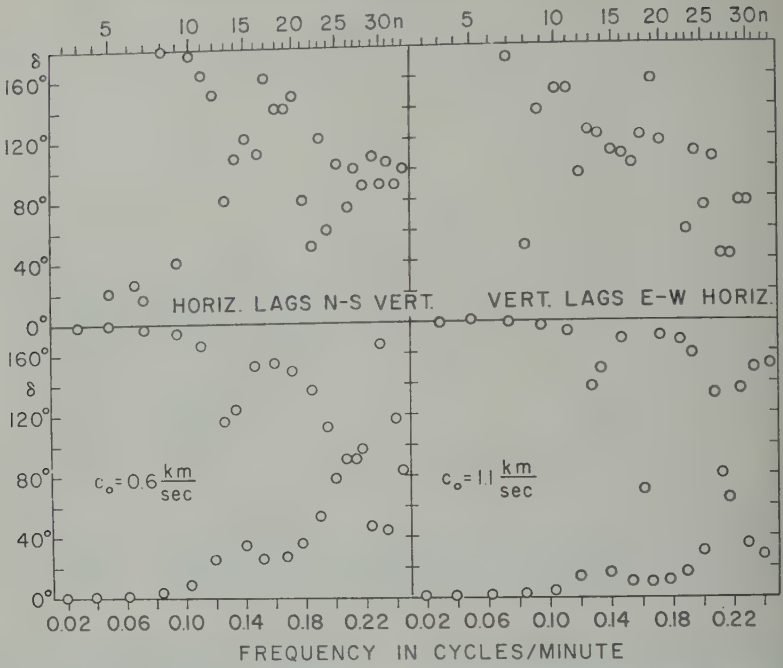


Fig. 6. (a) Experimental phase differences of Pasadena Z-NS and Z-EW seismographs for Chilean earthquake of 1960 for modes $n=5$ to $n=33$. (b) Theoretical phase differences for fault length of 160 km and rupture velocities of 0.6 and 1.1 km/sec.

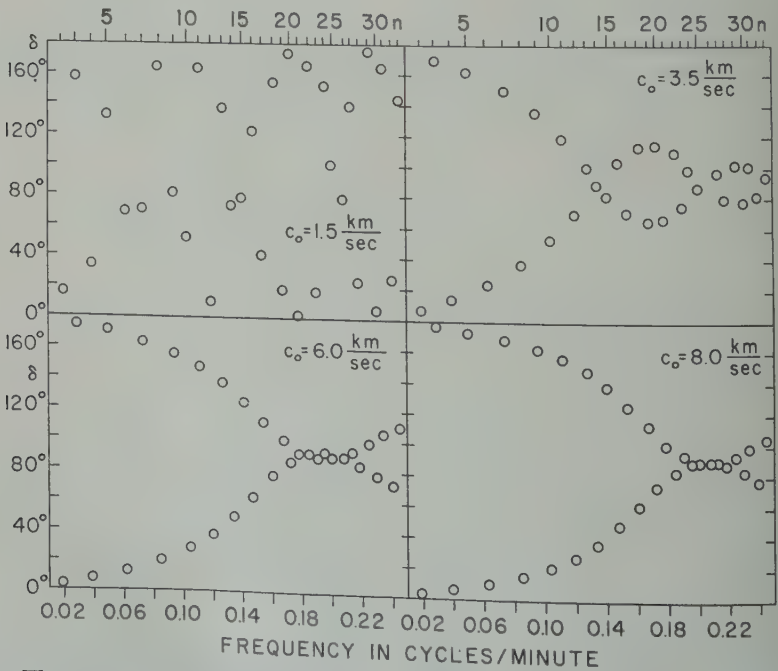
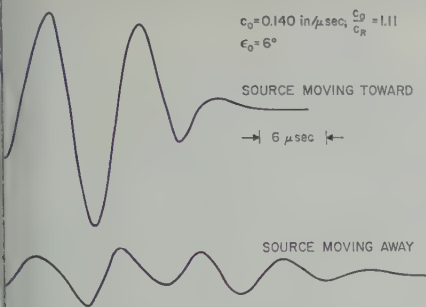


Fig. 7. Theoretical phase differences for fault length of 1100 km and rupture velocities of 1.5, 3.5, 6.0, and 8.0 km/sec.



Model seismograms of Rayleigh waves for moving toward and away from receiver.

cent. Since fault parameters are uncertain by several hundred per cent, the data may still be useful in the applications of this method. earthquake data. In applying this to the actual earth, we shall use the ratio of vertical and horizontal motions of motion observed in Pasadena

for the free oscillations of the earth excited by the great Chilean earthquake of May 22, 1960. A re-evaluation of the data presented by *Benioff, Press, and Smith* [1961] follows.

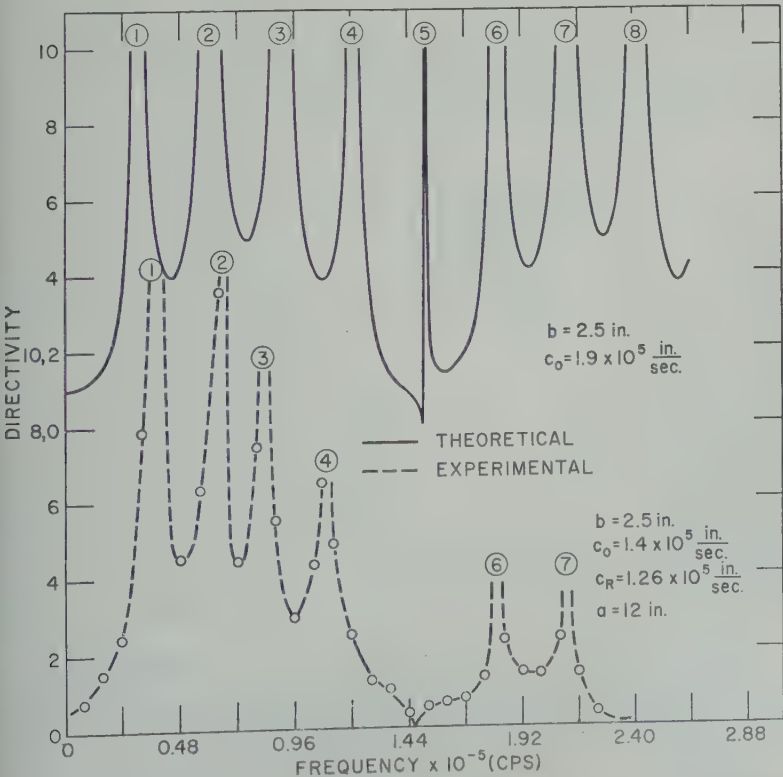
The theoretical expressions for phase shift as a function of fault parameters and mode number may be derived for a sphere using the traveling-disturbance concept referred to in the preceding section. The resultant expression for phase shift in the n th mode is

$$\delta_n = \frac{\pi}{2} + \bar{\varphi}_\theta^{(n)} - \bar{\varphi}_r^{(n)} \tag{7}$$

where

$$\tan \bar{\varphi}_r^{(n)} = \frac{\sum_{j=0}^n a_{(n-2j)} m_{(n-2j)}}{\sum_{j=0}^n a_{(n-2j)} l_{(n-2j)}}$$

$$S = \frac{n}{2}, \quad n \text{ even}; \quad S = \frac{n-1}{2}, \quad n \text{ odd}$$



9. Experimental and theoretical directivity functions for ultrasonic model experiment: fault length 2.5 inches.

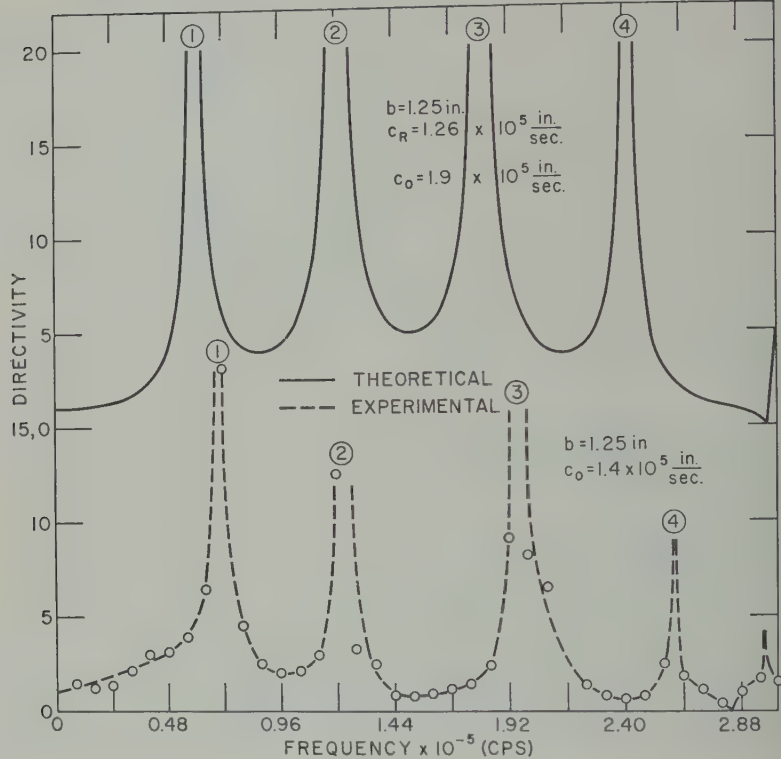


Fig. 10. Experimental and theoretical directivity functions for ultrasonic model experiment fault length 1.25 inches.

$$\tan \tilde{\varphi}_\theta^{(n)} = \frac{\sum_{j=0}^s (n-2j)a_{(n-2j)}P_{(n-2j)}}{\sum_{j=0}^s (n-2j)a_{(n-2j)}q_{(n-2j)}}$$

$$\begin{bmatrix} m_{n-2j} \\ P_{n-2j} \end{bmatrix}$$

$$= \frac{\sin Y_{n-2j}}{Y_{n-2j}} \sin \{(n-2j)\theta - Y_{n-2j}\} \\ \mp \frac{\sin X_{n-2j}}{X_{n-2j}} \sin \{(n-2j)\theta + X_{n-2j}\}$$

$$\begin{bmatrix} l_{n-2j} \\ q_{n-2j} \end{bmatrix}$$

$$= \frac{\sin Y_{n-2j}}{Y_{n-2j}} \cos \{(n-2j)\theta - Y_{n-2j}\} \\ \pm \frac{\sin X_{n-2j}}{X_{n-2j}} \cos \{(n-2j)\theta + X_{n-2j}\}$$

$$X_{n-2j} = \frac{\epsilon_0}{2} \left[\frac{\omega_n a}{c_0} - (n-2j) \cos \theta_0 \right]$$

$$Y_{n-2j} = \frac{\epsilon_0}{2} \left[\frac{\omega_n a}{c_0} + (n-2j) \cos \theta_0 \right]$$

In these expressions n is mode number,

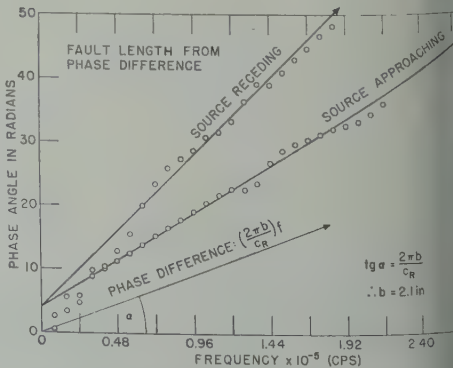
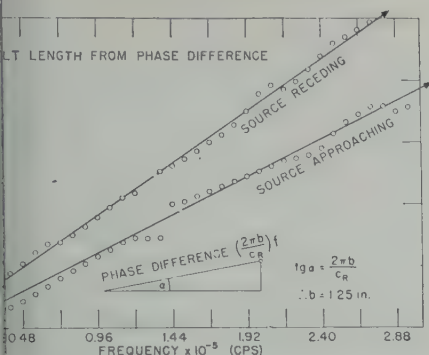


Fig. 11. Experimental differential phase for fault 2.5 inches.



Experimental differential phase for model fault 1.25 inches.

corresponding circular frequency, c_0 is source velocity, a_0 is angular fault length, θ is angular distance from epicenter to receiver, θ_0 is azimuth angle from epicenter measured from the direction of rupturing. Also, $a_{(n-2j)}$ is coefficient of the $\cos(n-2j)\theta$ term in the expansion of the zonal harmonic $P_n(\cos \theta)$. Experimental data [Benioff, Press, and Smith, 1961] were obtained from the Pasadena-component Press-Ewing seismograms of May 1960, $T_0 = 90$ sec) of the great Chilean earthquake of 1960. These were the longest-continued system available to us, and, as a result of the limited dynamic range of photorecords and the poor response to the higher modes, the amount and precision of the data were not satisfactory for full exploitation of the method. Record lengths of 1000 minutes were digitized at intervals of 0.2 minutes. The electronic computers 709 and 7090 were used to obtain phase differences from the cross-correlation of the vertical-NS and vertical-EW components. These are given in Figure 6a for station modes $n = 5$ to $n = 33$. Although the data show some scattering, trends are clearly visible. For example, the phase shifts for the lower modes are near 0° or 180° , as the theoretical ultrasonic experiments predict. From the 9th mode to the 15th or 16th mode the phase shifts rapidly converge towards the value of 90° . The phase shifts then diverge and converge again in the 29th or 30th mode in view of the approximations inherent in the theory and the roughness of the data, we should not expect to match each observed phase shift with a theoretical value. We should, how-

ever, be able to reach some conclusions about the faulting in the Chilean earthquake by matching the trends in the data just described.

Theoretical phase differences based on equation 7 were computed on the Seismological Laboratory's Bendix G-15D electronic digital computer. In Figure 6b results are presented for a fault length of 160 km and for rupture velocities of 0.6 km/sec and 1.1 km/sec. When compared with the observed values, it is seen that the theoretical values do not converge towards 90° as much as is required. Since smaller fault lengths or higher rupture velocities would increase this discrepancy, we conclude that the data are consistent with faulting in the Chilean earthquake which is much greater than 160 km.

Actually, evidence exists that the fault length was more like 1000 to 1200 km. Aftershocks following the earthquake extend over this length [Benioff, Press, and Smith, 1961]. Data from field intensities, sea level changes, and general geology also support a fault length of 1000 to 1200 km [St. Amand, 1961]. We have therefore assumed a fault length of 1100 km for the theoretical phase differences plotted in Figure 7. Various rupture velocities were assumed in an attempt to deduce the most probable value: 1.5 km/sec; 3.5 km/sec (shear velocity in crustal



Fig. 13. Disposition of recording stations and epicenter of Chilean earthquake of May 1960.

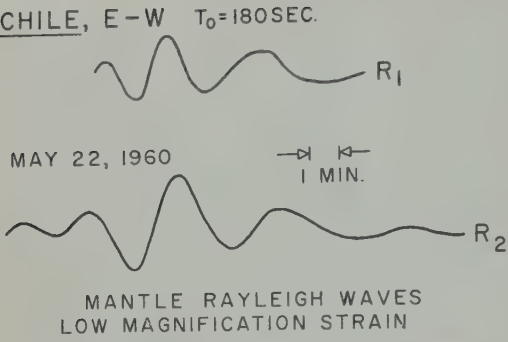


Fig. 14. Mantle Rayleigh waves R_1 and R_2 from the Chilean earthquake recorded on the Benioff strain seismograph in Pasadena, California.

rock); 6.0 km/sec and 8.0 km/sec (compressional velocity in crustal and mantle rock, respectively). Comparison with the observed data in Figure 6a suggests that a rupture velocity near 3.5 km/sec provides the best fit.

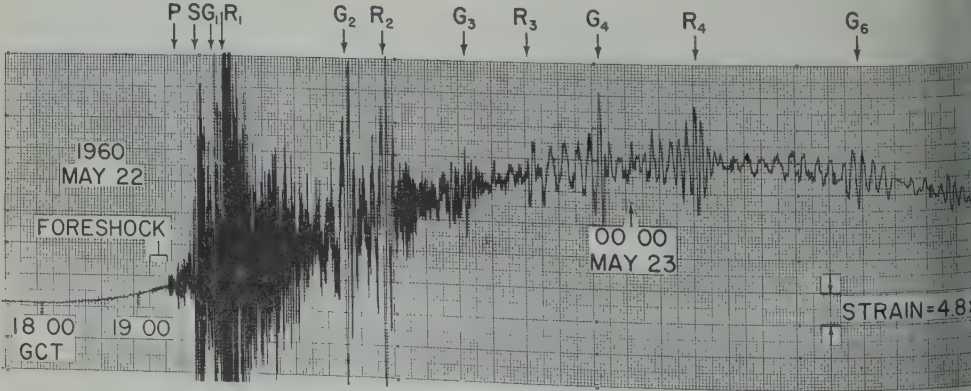
Fault parameters from mantle surface waves. Ben-Menahem [1961] has defined a directivity function equal to the ratio of spectral amplitudes of surface waves leaving a source in opposite directions. The convention is followed that the numerator of the ratio corresponds to waves leaving the source in the direction of rupture. This function is given by

$$D = \frac{\frac{c}{c_0} + \cos \theta_0}{\frac{c}{c_0} - \cos \theta_0} \frac{\sin \frac{\pi b}{\lambda} \left(\frac{c}{c_0} - \cos \theta_0 \right)}{\sin \frac{\pi b}{\lambda} \left(\frac{c}{c_0} + \cos \theta_0 \right)} \quad (8)$$

where c is the surface wave phase velocity corresponding to wavelength λ , c_0 is r velocity, b is horizontal fault extent, and the angle between the fault line and the circle from the epicenter to the seismic station. D is independent of the nature of source mechanism and assumes a constant rupture velocity over a finite horizontal distance. D is experimentally obtainable as a function of frequency by taking the ratio of the numerical Fourier transforms of surface waves which traveled the epicenter in opposite directions and the same distance. To obtain numerical transforms we can use even- and odd-order mantle surface waves recorded at the seismic station. The waves are equalized for differences in absorption due to path differences by deriving the absorption coefficients for even- and odd orders alone. No other equalization is required for path differences if we restrict ourselves to the long mantle waves recorded at the seismic station, since the geometric amplitude-distance factor is the same on a sphere and the displacement-amplitude-distance factor is removed by Fourier analysis. The factors c and γ are derived using even or odd surface waves [Sato, 1958].

By comparing experimental values of D with theoretical values computed from equation (8) we can, in principle, deduce the fault parameters b and c_0 .

Ben-Menahem [1961] has also suggested the use of the differential phase of mantle surface waves to recover fault length. The differential phase factors $\partial_1\varphi$ and $\partial_2\varphi$ are defined as follows



ISABELLA, CALIF. FUSED QUARTZ STRAIN SEISMOGRAPH.

Fig. 15. Benioff strain seismogram of the Chilean earthquake recorded at Isabella, California.

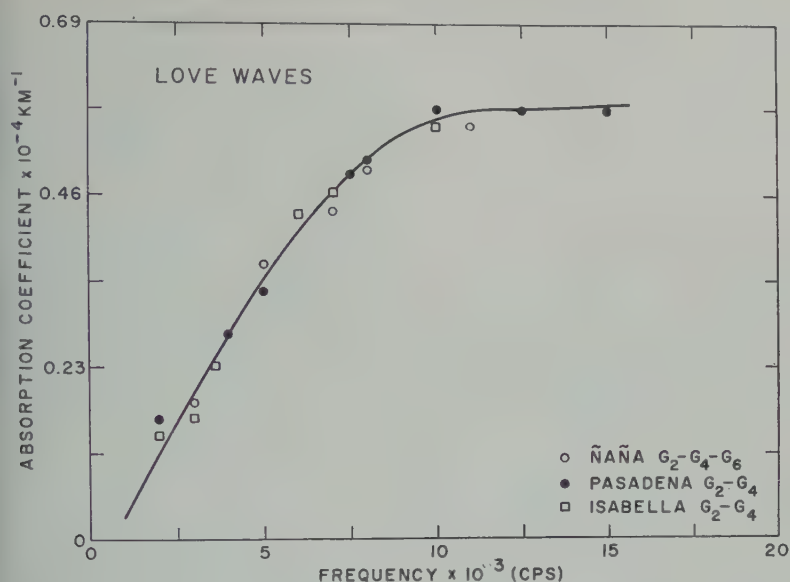


Fig. 16. Absorption coefficient for G waves from the Chilean earthquake.

$$\frac{-\varphi_n}{2\pi} = \frac{f}{c} (40,000 - 2 \Delta_1 + b \cos \theta_0) + M + \frac{1}{4} \quad (9)$$

$$M = 0, 1, 2, \dots$$

$$\frac{-\varphi_{n+1} - \varphi_n}{2\pi} + \left(1 - \frac{\Delta_1^\circ}{180^\circ}\right) (\varphi_{n+2} - \varphi_n) = \frac{fb}{c} \cos \theta_0 + \text{constant} \quad (10)$$

expressions φ_n is the phase of the Fourier transform of a mantle surface order n , having phase velocity c and f ; the shortest distance from epicenter is Δ_1 . The phase φ_n is taken with a common fiducial time.

ification occurs when phase velocity c and group velocity U are equal (no dispersion), U has a constant value U_0 . This last is for G waves in the period range 10-20 sec. For constant group velocity $U_0 = \eta_0$ where η_0 is a constant. If φ_n are redefined to be taken with respect to the running time of the wave, then

$$\frac{(fb/U_0) \cos \theta_0 + \text{constant}}{2\pi} \quad (11)$$

experimental procedure for applying the

differential phase is as follows: Obtain from digitized seismograms of surface waves the numerical Fourier transforms, extract the phases $\varphi_n, \varphi_{n+1}, \varphi_{n+2}$, as functions of frequency and form the quantities $\partial_1 \varphi$ and $\partial_2 \varphi$. Since f and c are known from auxiliary measurements, the factor $b \cos \theta_0$ can, in principle, be obtained.

Ultrasonic model study. The setup for modeling a faulting source was described earlier. It was used here with the modification that the

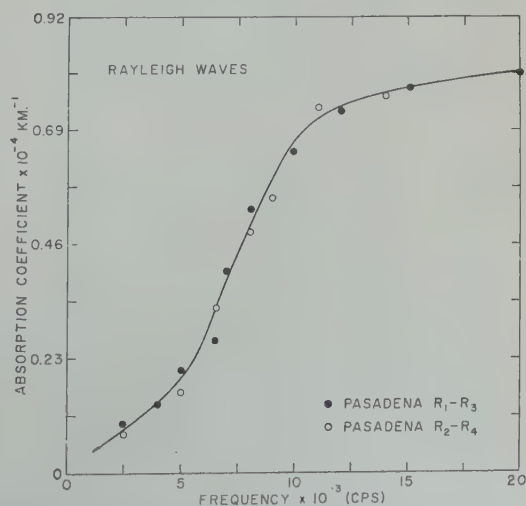


Fig. 17. Absorption coefficient for mantle Rayleigh waves from the Chilean earthquake.

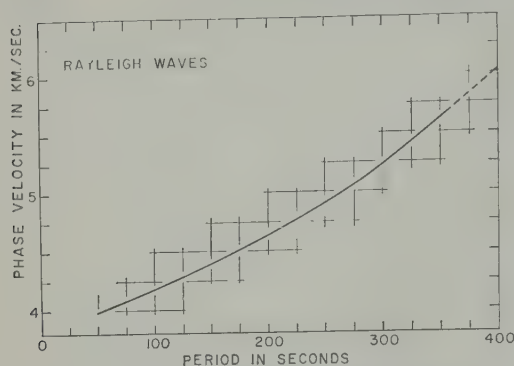


Fig. 18. Phase velocity of Rayleigh waves from the Chilean earthquake.

source emitted transient disturbances rather than sinusoidal waves. This change was made to simulate real conditions where surface waves are recorded as transients and are digitized and subjected to numerical Fourier analysis on an electronic computer.

Figure 8 shows model seismograms of Rayleigh wave radial motion at opposite sides of the source. Note the increased duration and the decreased amplitude of the waves emitted in the direction opposite to that of the source movement. In contrast, for the waves emitted in the direction of source movement the source transient is preserved and the amplitude is increased by constructive interference. These seismograms and others were digitized and Fourier analyzed to obtain the directivity functions and the differential phases.

Two cases were studied in which the fault lengths were 12° and 6° , the effective rupture velocity was 1.4×10^6 in/sec, and the recording distance was 120° . The ratio c_0/c_R was 1.11, and $\cos \theta_0 = \pm 1$ for the two-dimensional model.

The results for directivity are presented in Figures 9 and 10; differential phase results are given in Figures 11 and 12. We note that the directivity function has infinities, zeros, and minima. The infinities and zeros are given by $b(c_R/c_0 \pm \cos \theta_0) = n c_R/f$ according to whether the sign is plus or minus, and $n = 1, 2, 3 \dots$ is the order of interference. Thus, by locating the positions of several adjacent infinities and zeros, we can make a first estimate of b and c_0 . The entire function D is computed for various values of b and c_0 centering on the first estimates and compared with the experimental data to find those values of c_0 and b which provide the best

fit to the experimental data. It is seen in Figures 9 and 10 that a good fit is found to the experimental data when the theoretical fault length is the same as the experimental one and the theoretical rupture velocity is 1.9×10^6 in/sec, which is about one-third larger than the actual value.

Fault lengths derived from the differential phase are 2.1 and 1.3 inches (Figs. 11 and 12). These values are within 20 per cent of the actual values of 2.5 and 1.25 inches, respectively. The sensitivity of the method can be roughly gauged by noting that the peaks in Figures 9 and 10 fall at frequencies which are inversely proportional to $b(1 + c/c_0)$.

The ultrasonic model study indicates that the theory is essentially a correct one. Fault lengths are recovered with satisfactory reliability (20 per cent); fault velocities were recovered by 30 per cent. There are several sources of error which may account for this: (1) uncertainties in reading infinities and zeros in the experimental Fourier amplitude ratio curves; (2) errors introduced in digitizing the recorded data; (3) neglect of dispersion due to curvature of the model; (4) imperfect simulation of faulting in the model; (5) a theoretical error of the order of b/Δ . This amounts to 10 per cent in the experiment for a fault length of 2.5 inches.

Chilean earthquake data. Mantle Rayleigh waves and G waves excited by this great earthquake circled the earth many times—in the direction opposite to the direction of faulting. We use the data from the stations at Pasadena and Isabella shown in Figure 13. Distances and values of θ_0 deduced from field evidence are also shown in Figure 13. Field data [St. Amant, 1961] and the large amplitudes of even-order

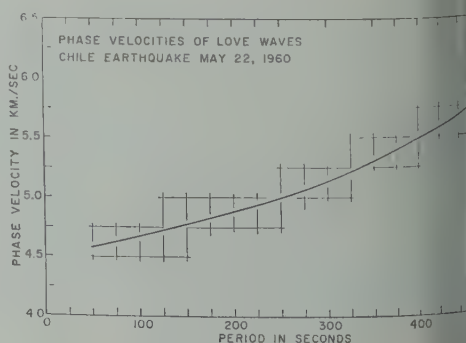
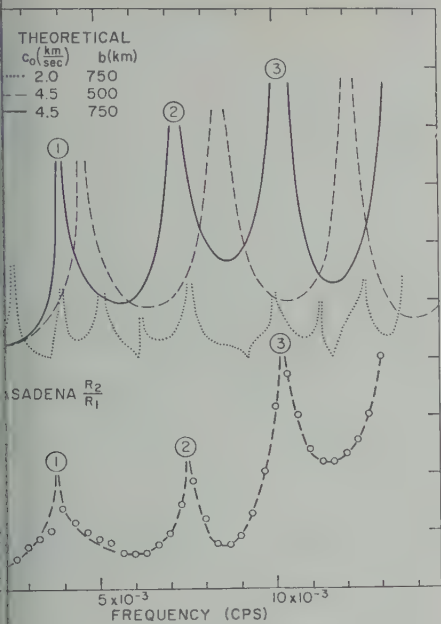


Fig. 19. Phase velocity of Love waves from the Chilean earthquake.



20. Experimental and theoretical directivity functions for Pasadena R_2/R_1 from Chilean earthquake.

and G waves compared with odd-order waves leave no doubt that the rupture started and progressed with finite velocity in the direction of the arrow. St. Amand writes (personal communication, 1961): 'The fault runs to the coast about 20-50 km offshore at a distance of 1000-1200 km. The rupture did not slip at the south end of the Arauco Peninsula and the rupture propagated southward just north of the Taitao Peninsula.' Another source of information about fault length and rupture velocity is the remarkable duration of 6 or 7 minutes for the T phase in Hawaii [Eaton, Richter, and Ault, 1961]. This duration is consistent with a fault length of 1200 km and a rupture velocity of 3 to 4 km/sec.

Parameters of the seismograms used are given in Figures 14 and 15. The numerical procedure for digitization (intervals 12-16 sec), followed by Fourier analysis on an IBM 709 computer. Rayleigh waves were windowed with fixed group velocities of 3.90 and 3.47 km/sec. Most G waves fell in the group velocity range 3.8 to 4.44 km/sec. The numerical data used to compute absorption, phase velocity, and differential phase came from

the Fourier analysis. The absorption coefficients are shown in Figures 16 and 17. The formula used for obtaining phase velocity is

$$C(T) = \frac{40,000}{\partial t - T(\partial\varphi + M - \frac{1}{2})} \text{ km/sec} \quad (12)$$

where $\partial\varphi$ is the phase difference between surface waves of order n and $n + 2$, ∂t is the difference in their arrival time, M is an integer, and the factor $\frac{1}{2}$ comes from the polar phase shift. M was fixed by requiring agreement of specific phase velocities with values derived from the free oscillations of the earth. The Rayleigh wave results in Figure 18 agree reasonably well with the values given by Brune, Nafe, and Alsop [1961]. The G (Love) wave data in Figure 19 are considered sufficiently accurate for our purposes, but their precision may not be good enough for mantle structure investigations. Precision may be lost with G waves because the flat group velocity curve gives the waves the character of transients whose beginning and ends are difficult to define.

As described earlier, the ratios of the Fourier transforms of successive orders of mantle surface waves (multiplied by a correction factor to equalize for absorption because of path differences) yield the directivity function. For example, for the Pasadena ratio R_2/R_1 the path difference is 21,542 km and the correction factor is $e^{21,542 \gamma_R}$ where γ_R is obtained for the appropriate frequency from Figure 17.

Observed and theoretical directivity functions for Pasadena R_2/R_1 and Isabella G_4/G_3 are presented in Figures 20 and 21. The theoretical values are computed from equation 8. Several

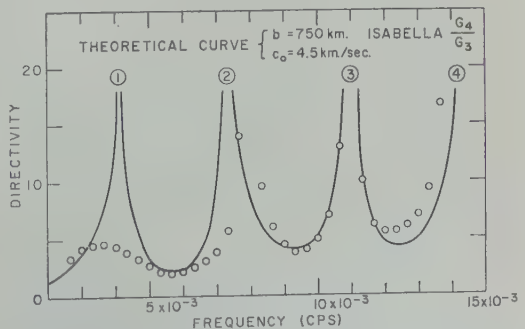


Fig. 21. Experimental and theoretical directivity functions for Isabella G_4/G_3 from Chilean earthquake.

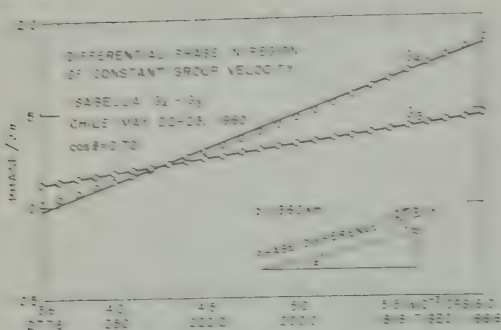


Fig. 22. Differential phase data for Isabella $G_1 - G_2$ from Chilean earthquake.

theoretical curves are given for the Rayleigh wave interpretation to show the sensitivity of the method. A fault length of 750 km and a rupture velocity of 4.5 km/sec is seen to provide reasonably good fit to the data. That the observed data show a succession of maxima and minima as theoretically predicted gives us confidence in the general validity of the method. In fact, as predicted by the theory, it has been found that the Fourier transforms of the odd-order mantle R and G waves all show succession of minima, while the even-order transforms are comparatively smooth.

Data for obtaining the differential phase between G_1 and G_2 recorded at Isabella are given in Figure 22. Using equation 11 and Figure 22, we find $\partial\varphi = 220^\circ$ or $b \cos \theta_0 = 220 \times 4.38$. Using $\theta_0 = 45^\circ$ we obtain $b \sim 1360$ km.

The experimental data for the Chilean earthquake leave much to be desired. However, the approximate agreement with the free-oscillation results and similarities with the ultrasonic experiments indicate to us that we are observing real phenomena which serve to indicate roughly the fault parameter.

Discussion. We have described three methods of obtaining fault length and rupture velocity. The methods are not independent because they have the same physical basis, namely the radiation pattern of a moving source. To the extent that different data are used, the methods are supplementary.

Ultrasonic experiments serve to demonstrate the general validity of the methods. Discrepancies were found which typically were about 20 or 30 per cent. For the Chilean earthquake the methods indicated fault lengths in the 800 to

1400-km range and rupture velocities close to shear waves in crustal rock than to compressional waves. In view of the fact that these parameters were previously uncertain by several hundred per cent, we consider these results to be of some importance. Despite the uncertain analogy to faulting, it is interesting to note that rupture velocities corresponding to Rayleigh velocity (which is within 10 per cent of the shear velocity) are predicted theoretically for a moving Griffith crack (Yoffe, 1951).

Many oversimplified assumptions are made in the theory which can be corrected in a future development. Variations in rupture velocity, source strength and allowance for bending of the faulting will have to be made. This will introduce additional fault parameters which may be difficult to determine separately unless more data are available or unless auxiliary information is used. Instrumental developments now in progress will provide more complete data in the future.

The finiteness of the seismic source can affect the value of the group velocity as measured from the records of a single station by the conventional method. To compute this effect we regard the wave form as being expressible as a Fourier integral with the phase factor

$$\varphi = (\omega t - k \Delta) - (kb/2)(c/c_0 - \cos \theta_0) \\ 0 \leq \theta_0 \leq 2\pi$$

Evaluating the wave-form integral by the method of the stationary phase, we find

$$U' = \frac{U}{1 + \frac{b}{2\Delta} \left(\frac{U}{c_0} - \cos \theta_0 \right)} \\ \approx U \left[1 - \frac{b}{2\Delta} \left(\frac{U}{c_0} - \cos \theta_0 \right) \right]$$

where $U = dx/dk$ and $U' = \Delta/t$ is the group velocity as measured from the seismic records. Equation 14 thus states that U' is the group velocity with allowance for the fault length and rupture velocity (the group velocity without allowance for the fault) except for the case $|\theta_0| \leq \pi/2$ and $U/c_0 < \cos \theta_0$. In other words, the finiteness of the source will modify the group velocity curve and make it dependent on the source parameters and the position

with respect to the fault. To equalize moments, (14) will have to be solved three 'corrected' for the source. The two source functions on the phase depends to a certain degree on the way used, and it will not be dealt with here. Source effect for a 1000-km fault can be as 10 per cent. Previously, about 1 was thought to be the capability of velocity method. Use of group velocities before be restricted to higher-order phase when dealing with large faults.

gments. The research was partially by contract AF-49 (608) 910 of the Air Force Applications Center as part of and Research Projects Agency project

ly acknowledge use of the IBM 709 facilities at the Western Data Processing C.L.A., and the IBM 7090 facilities at Radiation Laboratory, California Institute of Technology.

REFERENCES

- Further study of the mechanism of oceanic earthquakes from Rayleigh waves, *J. Geophys. Res.*, **65**, 4165-4172, 1960.
- F. Press, and S. Smith, Estimation of motions of the earth by earthquakes, *Am. J. Science*, **259**, 695-718, 1961.
- H. A. Radiation of seismic surface waves from moving sources, *Bull. Seism. Soc. Am.*, in press, 1961.
- Brace, J. J. Nale, and J. Oliver, A simplified method for analysis and synthesis of dispersed surface wave trains, *J. Geophys. Research*, **65**, 287-304, 1960.
- Brace, J. J. Nale, and L. Akop, Polar phase shift of surface waves on a sphere, *Bull. Seism. Soc. Am.*, **51**, 247-257, 1961.
- Eaton, J., D. H. Richter, and W. U. Ault, The tsunami of May 23, 1960, on the Island of Hawaii, *Bull. Seism. Soc. Am.*, **51**, 135-157, 1961.
- Fung, M. W., S. J. J. and F. Press, *Elastic Waves in Layered Media*, McGraw-Hill Book Co., New York, 1957.
- Healy, J., and F. Press, Two-dimensional seismic models with continuously variable velocity depth and density functions, *Geophysics*, **25**, 947-957, 1960.
- Hodgson, J. H., Editor, The mechanics of faulting (a symposium), *Publ. Dom. Observ., Ottawa*, **20**, 252-418, 1957.
- Knapoff, L., and F. Oliver, Radiation from a strike-slip fault, *Bull. Seism. Soc. Am.*, **43**, 163-178, 1953.
- Oliver, J., Rayleigh waves on a cylindrical curved surface, *Earthquake Notes*, **26**, 24-25, 1955.
- Satô, Y., Attenuation, dispersion and the waveguide of the G wave, *Bull. Seism. Soc. Am.*, **48**, 231-251, 1958.
- St. Amant, P., Algunas observaciones y explicaciones acerca de los terremotos Chilenos de 1960, *Publ. No. 2 de la Escuela de Geología de la Universidad de Chile*, in press, 1961.
- Yaffé, E. H., The moving Griffith crack, *Phil. Mag.*, **42**, 739-750, 1951.

(Manuscript received June 22, 1961.)

SH Motion from Explosions in Soil

C. KISSLINGER, E. J. MATEKER, JR., AND T. V. McEVILLY

*Department of Geophysics and Geophysical Engineering
St. Louis University, St. Louis, Missouri*

Abstract. Although not predicted by simple theory, prominent tangential horizontal motion observed in almost all cases from both chemical and nuclear explosions. A series of experiments designed to clarify the factors that are significant in producing this motion has been carried out. The seismograms, recorded at small distances, clearly indicate the asymmetrical pattern of radiation of the *SH* motion in the source region, as contrasted with the symmetrical pattern for the radial and vertical motion. This radiation pattern is the most useful indicator of the nature of the generating mechanism. The combined effect of charge size and depth determine which of two wave forms will appear. The most prominent *SH* motion is in the form of Love waves, but this motion is much more pronounced for shallow shots that produce cratering, or at least surface cracking, than for completely contained shots. An examination of theoretical radiation patterns indicates that crack formation may contribute significantly to the generation of *SH* motion.

Introduction. Explosions produce prominent tangential transverse, or *SH*, motion. Hundreds of three-component records of explosion-generated ground motion made under widely varying circumstances testify to this. This motion is observed whether the source is a chemical or a nuclear explosion. With regard to *SH* motion, see, for example, [1959], in the Berkner panel report. In simple models which approximate the actual theoretical analyses show that for explosions the shear waves will be polarized in the plane of propagation . . . In such simple models the existence of shear waves with a transverse horizontal component (*SH*) would be an unambiguous indication of a non-explosive source. This theoretical result, unfortunately, does not apply to the real earth, where transverse horizontal components of motion are customarily recorded from explosions. Several mechanisms can be imagined which could cause this effect, the simplest of which is the existence of geological discontinuities which are transverse horizontal. It should be noted that the oversimplification of the mathematical model of the earth is properly pointed out, no conclusion is made of the possibility of an overestimation of the model of the explosion as a source of seismic waves. Leet [1946], in his presentation of the first publication in the open literature of the ground motion from a nuclear

explosion, pointed out that the assumption that an explosion is a radially symmetric simple expansion about a point does not agree with observation. Probably our willingness to accept as a model of a contained explosion a symmetric distribution of normal stress on the walls of a spherical equivalent cavity stems from the excellent agreement between observation and Sharpe's theory for *P* waves [Sharpe, 1942]. Since the principal interest in explosion-generated waves, at least until recently, has been in seismic prospecting and the *P* wave, perhaps there has seemed no point in investigating a more elaborate model. Indeed, this mechanism seems to be adequate to explain the properties of compressional waves, but this does not rule out the possibility that a more complex action takes place in the neighborhood of the explosion which results in the direct production of shear waves with arbitrary polarization.

The experimental program. Since August 1960, an intensive experimental investigation of the nature and properties of explosion-generated *SH* motion has been under way. Although experiments have been carried out for several media, this report deals only with the results for explosions in soil consisting of about 100 feet of loess and Pleistocene clay, overlying Mississippian limestone. Charge sizes have been varied from $\frac{1}{4}$ to 2 pounds of dynamite. Most ob-

servations have been made at distances from 15 to 30 meters, with two profiles extending to 195 meters from the source. The primary emphasis has been on the observation of surface displacements in the neighborhood of the source, so that effects of intervening geology are suppressed relative to the effects of controllable parameters at the source. All observations have been made with conventional three-component portable seismographs, modified for remote operation and including a firing-time signal. The instruments were precisely aligned by means of a transit set up over the shot point.

Although long profiles have been shot to obtain travel-time and dispersion data, the most useful observations have been obtained by surrounding the source with instruments at equal distances and observing variations in wave form, amplitude, and frequency, with azimuth. The approach is much the same as that of the earthquake seismologist who attempts to reconstruct the focal mechanism from the distribution of

amplitudes over the earth's surface. It is that some of the opinions as to the origin of *SH* motion are colored by the lack of observations in many directions from the source. So it seemed prudent to avoid possible unwarranted assumptions as to the important factors governing *SH* generation, the approach has been to observe the radiation pattern for a variety of source conditions and from these patterns seek a mechanism of generation.

Of the 29 shots that have been fired at the test site, the data from only a selected group are discussed here. These data are concerned with the basic wave form and the effect of charge size, depth, and degree of containment. The basic character of the wave form is seen in Figure 1. The upper record was obtained at 25 meters from 1 pound of dynamite tightly tamped at a depth of 0.8 meter; the center at 105 meters from 1 pound, tightly tamped at 1.3 meters and the lower at 165 meters from 1.5 pounds tightly tamped at 1.2 meters depth. The traces are vertical, radial, and tangential motion, from top down, timing lines are .02 sec apart and the firing time is on the fourth trace at the bottom. The vertical magnification is about twice that of the horizontals. The radial and vertical motion at this site has been discussed in previous work [Kisslinger, 1959].

The transverse trace on the upper record is typical of those to be discussed. A small-amplitude train of high-frequency (25 cps) motion precedes a larger-amplitude event at about 0.8 sec. At 15 meters these same events were found to have frequencies of 30 to 50 cps and 15 cps respectively. At the larger distances the high-frequency events, probably body *SH* waves, are strongly attenuated, while the longer-period events, now showing marked dispersion, have at least maintained their amplitude relative to the Rayleigh waves. The time of onset of this dispersed train corresponds to a surface velocity of about 175 m/sec.

The statement that these are typical transverse wave forms is not to be interpreted to mean that all records show motion of this form. Indeed, the transverse motion, at least at short ranges, varies markedly in character in different directions from one shot and from shot to shot. Data on the azimuthal variation at large ranges have not yet been collected. However, the motion seen here has been observed often enough

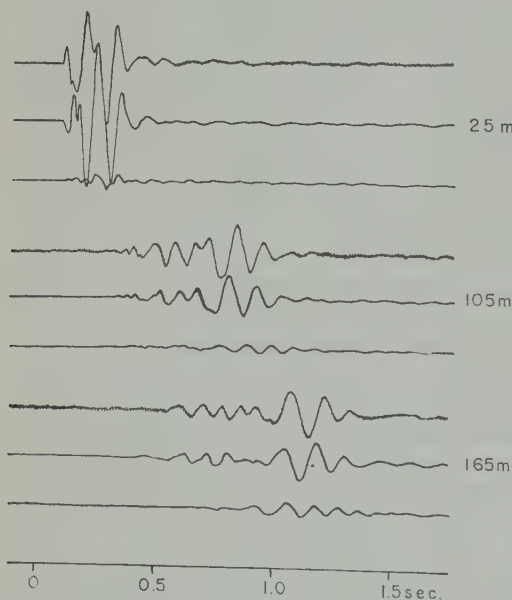


Fig. 1. Three-component records at Florissant site, showing variation of wave form with distance. The traces are vertical, longitudinal, and transverse (top to bottom) on each record. All shots tightly tamped. Upper record, 25 meters from 1 pound at 0.8 meter depth; center record, 105 meters from 1 pound at 1.3 meters depth; lower record, 165 meters from 1½ pounds at 1.3 meters depth.

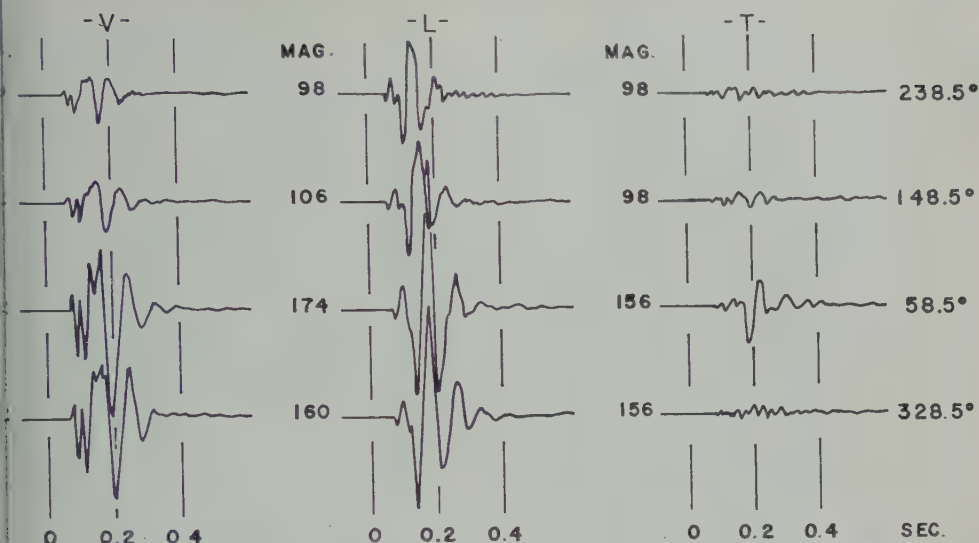


Fig. 2. Three-component records at Florissant site, showing variation of wave form with azimuth at a fixed distance of 15 meters. Static magnifications are given at left of each trace, azimuths from north at right.

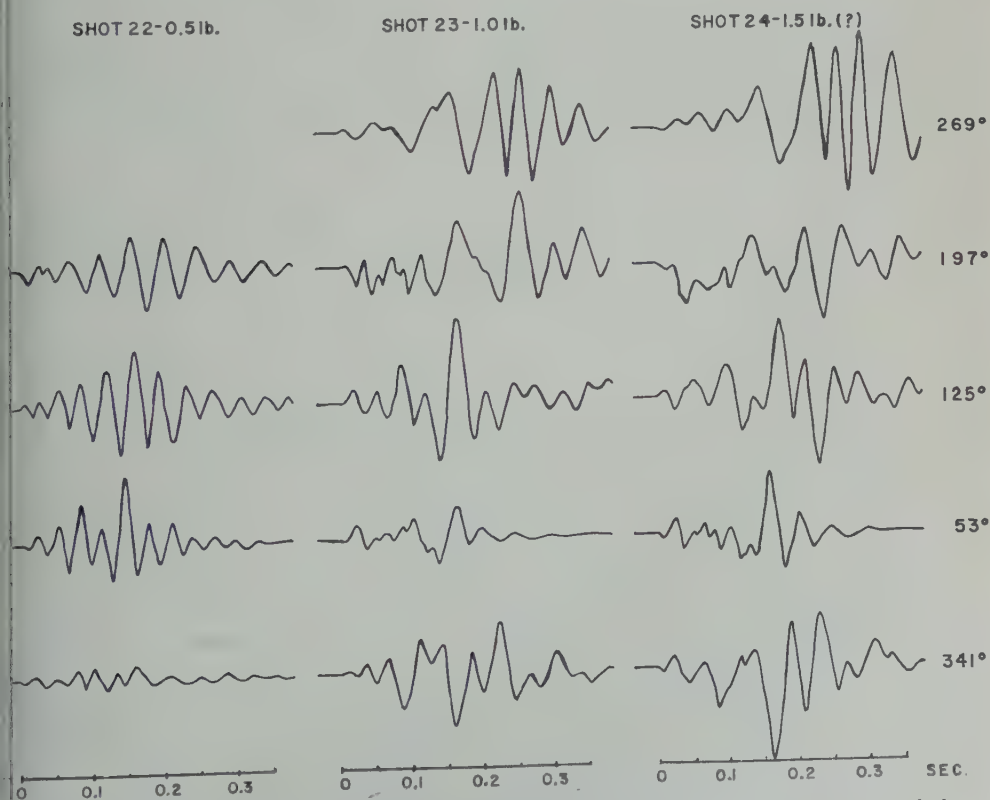


Fig. 3. Effect of charge size on SH motion. All shots at a depth of about 1.4 meters. Azimuthal array, r approximately 25 meters.

to be considered the basic wave form, of which others are variations.

Because the lack of symmetry of the transverse motion is fundamental to this study, it is important to note that the radial and vertical components show remarkable reproducibility, even in fine detail, for all records of a single shot, and from different shots, where depth of source and detector distance are unchanged. Therefore, the lack of symmetry of *SH* motion, as well as the departures from reproducibility, must be taken as real phenomena, indicative of the generating mechanism. Also, the independence of the transverse component from the other two is apparent. This last point is significant because it indicates that there is little value in drawing particle motion diagrams which combine the transverse trace with either of the others, at least for the surface wave portion. The usual procedure of treating the *SH* component separately from the radial and vertical in theoretical work [Ewing, Jardetzky, and Press, 1957, p. 25] finds full support in these observations.

Records and radiation patterns. The three components of motion recorded at four azimuths, 15 meters from 0.5 pound at 1.3 meters depth, are shown in Figure 2. These are tracings of the original records. The differences in mag-

nifications should be noted. When these instrumental differences are taken into account, the symmetry of the vertical and radial motion is remarkable, and the lack of symmetry of transverse motion is equally striking. All motions are with respect to magnetic north.

As a test of the effect of charge size, three shots were fired at a depth of about 1.5 meters. Shots of 0.5, 1, and 1.5 pounds were recorded at approximately 25 meters on five instruments spaced around a circle. The three shot patterns were 3 meters apart on a line. The recording sites were the same for all shots. The tangential ground motion is shown in Figure 3. These are actual ground amplitudes, plotted with the observable motion at zero time. A distinct difference between the character of the motion from the small shot and that for the two larger shots is seen. The blank space for shot 24 is due to an instrument malfunction. There is a possibility that the largest charge was not to full strength, perhaps owing to a defective stick of dynamite. However, shot 24 did produce noticeably more surface cracking than shot 23. The data indicate that, for this depth, there is a charge size between 0.5 and 1.0 pounds at which the wave character changes.

Although the same sequence of events can

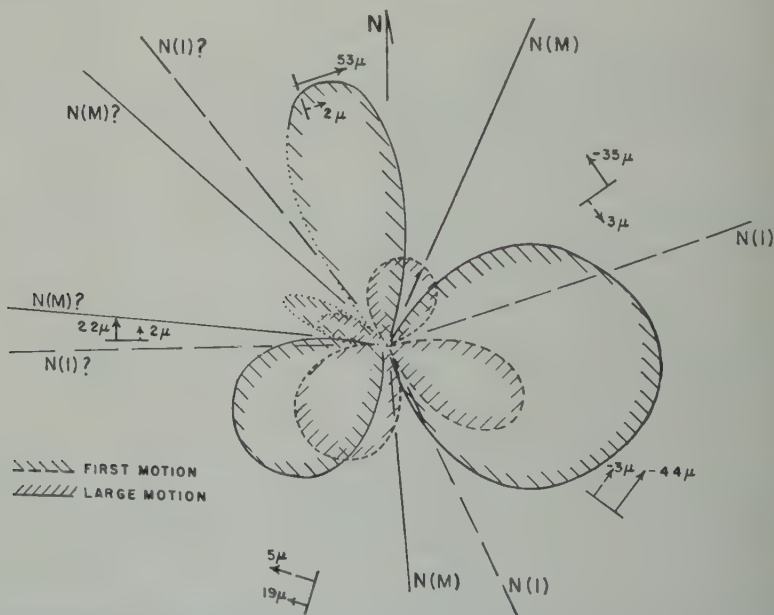
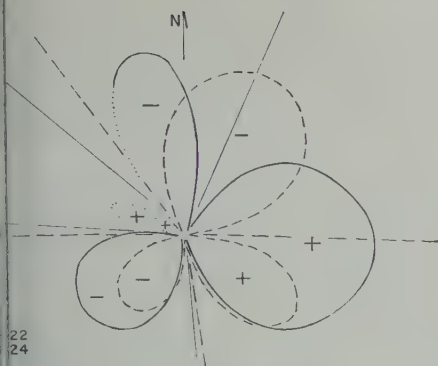
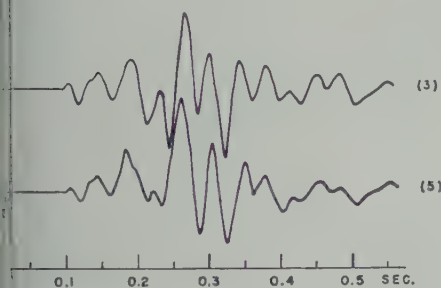


Fig. 4. Radiation of *SH* motion, shot 24. Patterns based on first motion and large motion are compared.



Comparison of radiation patterns of SH motion from cratering shot (24) and contained shot (22).



Reversal of polarity of entire SH wave at two stations 144° apart (shot 24).

the three records for a given instrument. The small shot yielded an apparent beat-form with almost constant frequency throughout. The succession of small-amplitude, high-frequency motion, followed by large-amplitude, lower-frequency motion, is apparent for larger charges. The early high-frequency motion (first tenth of a second) is relatively constant and tends to blend continuously with the motion for the smallest shot. It will be observed that the effect observed here is a function of charge size alone, but depends on degree of containment. The variation of form with azimuth is obvious for each of the three.

The following technique was employed to reduce the nonsymmetry to a graphical form. The peaks and troughs were correlated around the world. The correlation was based on travel time of the wave crests. The correlation was not perfect in some cases, especially for an instrument that happened to be near a node of SH

motion. The amplitude and direction of a few events were noted and plotted to scale on polar coordinate paper at the correct angular position. First motion and one or two peaks in the maximum amplitude portion of the record were used. Nodes were located by interpolation between positions at which the direction of motion was reversed. In general, the positions of nodes based on first motions agreed with those determined from later events. The radiation pattern was then drawn by plotting the amplitudes to scale along the radial lines and drawing in smooth lobes consistent with the nodal lines. It should be emphasized here that the radiation patterns in some cases are quite complex; with only five instruments it is doubtful that all the details have been observed. This is especially true because the pattern occurs more or less at random relative to the instrument layout, and the presence of an instrument near a node or an antinode is coincidental. It turns out to be quite possible for an entire lobe to fall between two instruments, in which case it can only be sketched in by inference.

The radiation pattern for shot 24, the 1.5 pound shot, is shown in Figure 4. The dashed curves are for first motion, the solid for the large motion. The positions of the nodes are marked around the edges of the figure, with $N(1)$ referring to a node based on first motion and $N(M)$ to one based on the large-amplitude motion. The southern nodes agree well, being 20° apart. However, the northern lobes are rotated so that the maximum from first motion falls on the northeastern node of large motion. This array was such as to leave doubt as to what happens in the NW quadrant. The figure was not the first one drawn, and the missing quadrant, indicated by the dotted curves, was filled in on the basis of prior observations. The amplitudes indicated near the arrows are in microns.

Figure 5 shows the superimposed large-motion radiation patterns from shot 22 (0.5 pound) and shot 24 (1.5 pounds). The patterns do not agree closely, the northern lobes again being displaced. The pattern from the small shot (22) agrees very well with that for the first motion of the large shot (24, Figure 4). This indicates that the predominant SH motion from the small, completely contained shot is generated by a mechanism similar to that generating the

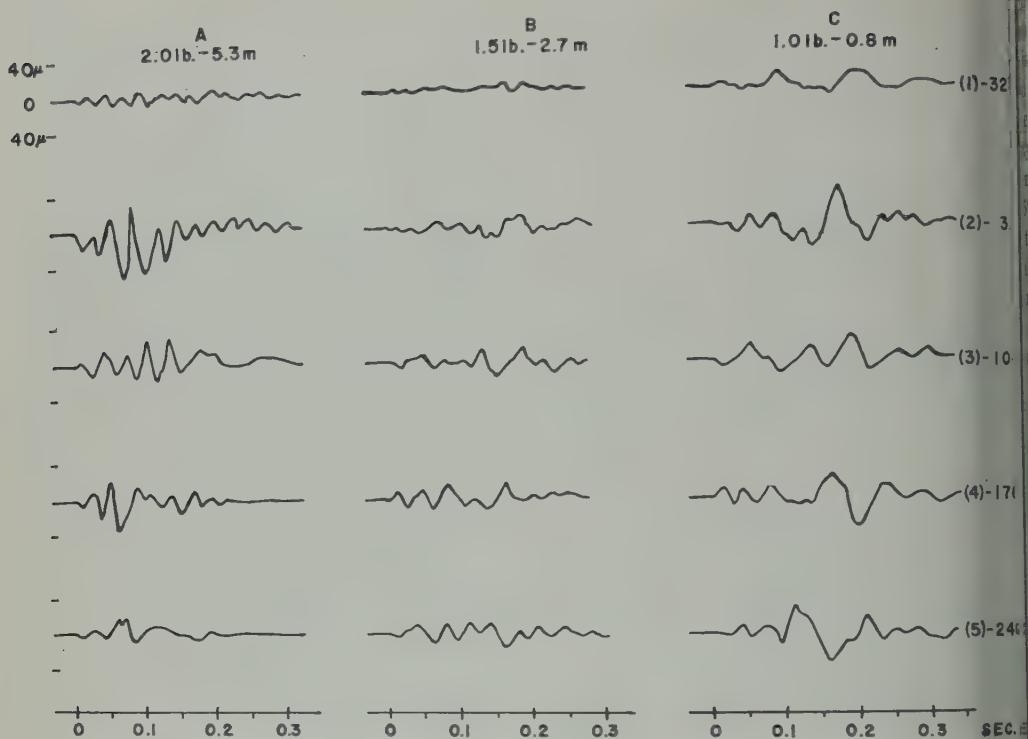


Fig. 7. Effect of charge depth and size on SH motion. Azimuthal array, $r = 25$ meters.

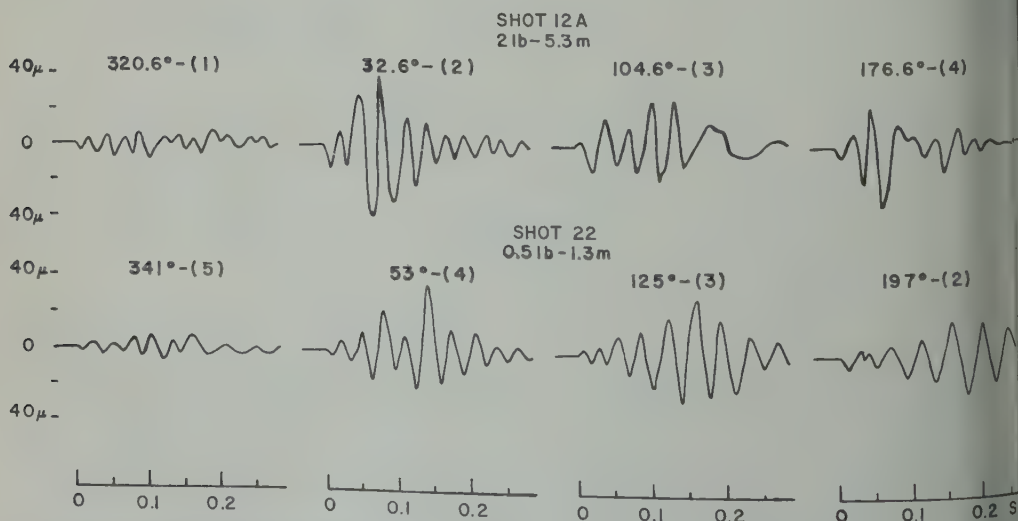
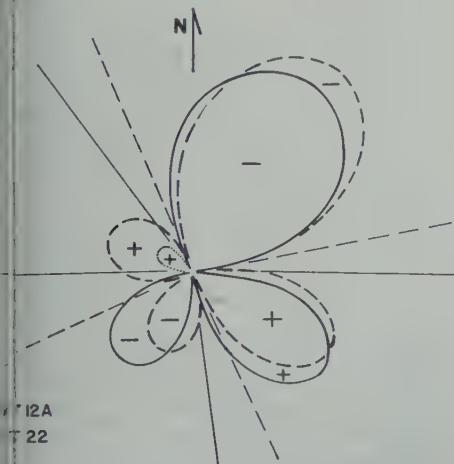


Fig. 8. Comparison of wave forms at approximately equal azimuths from similarly contained shots of different sizes. Shot points approximately 20 meters apart.

tion for the large shot. There is some
e, if not in the actual mechanism at
its orientation, for the longer-period
predominating in the second case.

emphasize the reality of the reversal of
of motion at corresponding points on
lobes, the traces from the instruments
and 341° for shot point 24, with the
face inverted, are reproduced in Figure 6.
most perfect agreement indicates that
the wave form, in its entirety, is propa-
with reversed polarity in roughly oppo-
sitions. This has been observed for many

7 shows the transverse ground mo-
5 meters from three shots in the same
ounds at 5.3 meters, 1.5 pounds at 2.7
and 1 pound at 0.8 meter. The deep shot,
is completely contained; the shallow
7, created a distinct crater zone. The
node at 321° and maxima near 33° and
in be seen in the column for any one
the change in character of motion for
shots is seen along the rows of the
the deep shot, largest in charge size,
predominantly high frequency motion
duration. The shallow shot again pro-
the sequence of small, high-frequency
followed by the large-amplitude, longer-
motion. The intermediate shot yielded
between these two; apparently the
nod motion was not well developed, and
er frequency was reduced. This gives an
on of a much smaller motion, even



Comparison of radiation patterns of SH motion from two contained shots.

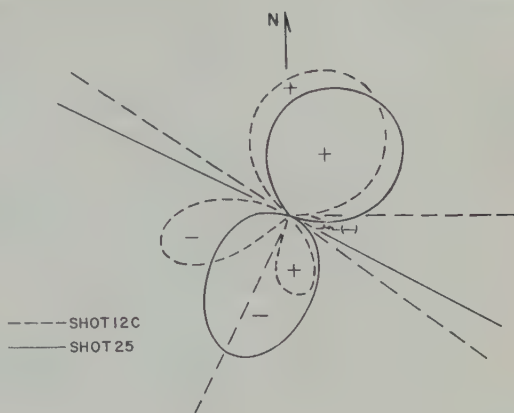


Fig. 10. Comparison of radiation patterns of SH motion from two cratering shots.

though the total energy release is greater than for the shallow shot. On the basis of these data, as well as the dispersed wave train of Figure 1, it is concluded that the longer-period motion is a true surface wave, a Love wave, and is reduced in amplitude as the depth of the source is increased. It is worth noting that the peak-to-peak vertical Rayleigh wave amplitude was 190μ for the 1-pound shallow shot and about 60μ for the 2-pound deep shot.

The previous data on varying charge size at constant depth indicate that depth alone cannot be the primary factor in determining which of the two wave forms will appear, but that the combination of depth and charge size as revealed by the containment of the shot is paramount.

To emphasize this point, the transverse motions from the two contained shots, shot points 12A and 22, are compared in Figure 8. It is apparent that, in spite of the differences between 2 pounds at 5.3 meters and 0.5 pound at 1.3 meters, these records are more like each other than they are like the other records with which they have been compared above. Note that the records paired here were not made by the same instruments (instrument numbers are in parentheses following the azimuth). The surprising result is that the radiation patterns prove to have the same orientation with respect to geographic directions. Not only is it possible to find a record from each shot that is similar to one from the other, but the radiation pattern seems to be tied into the geography of the test site.

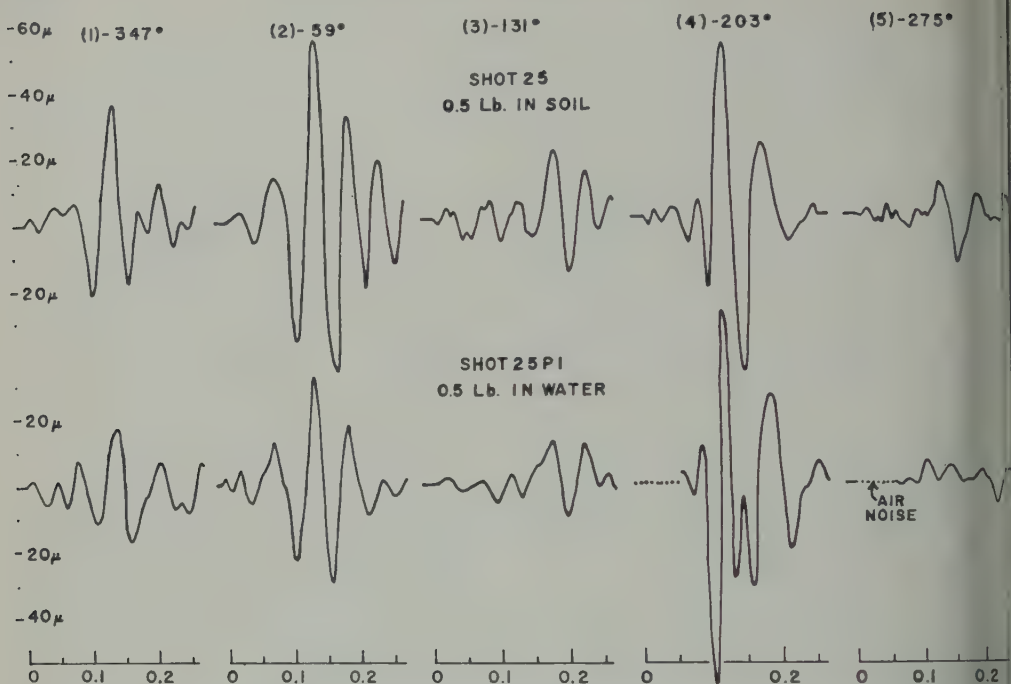


Fig. 11. Comparison of wave forms showing effect on *SH* motion of replacing soil around shot with water.

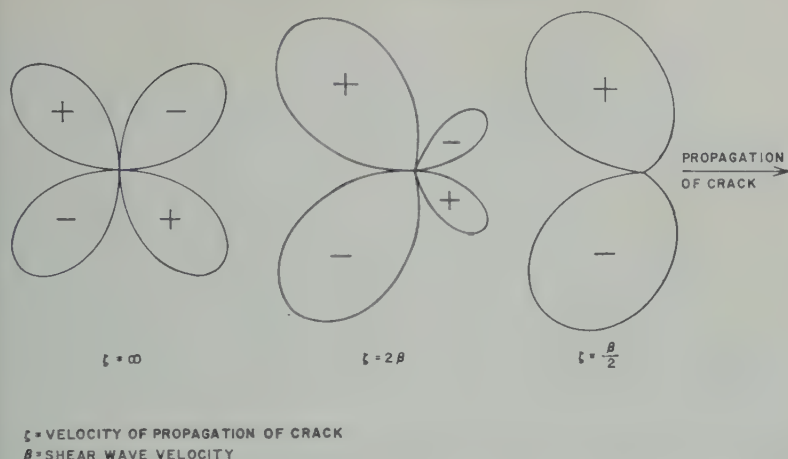
The superimposed polar plots of these two shots are shown in Figure 9. The similarity of the patterns bears out the previous statement. The discovery of a preferred orientation for the radiation pattern was unexpected, but it is verified by other data over a test area of about 10,000 m².

Figure 10 shows a similar comparison of polar plots for shots 12C and 25, both of which produced well-developed craters. Shot 25 was 0.5 pound at 0.2 meter. While the similarity of actual wave form is not as clear as in the case of the contained shots, the over-all character is similar, and the locations of nodes and reversals of polarity are also similar. Shot 25 produced one of the few two-lobed patterns observed. The larger and deeper cratering shot has produced a more complex pattern, with splitting of the southern lobe.

The final experiment to be described was performed by digging out the crater formed by shot 25 to make a smooth-walled hemispherical pit, larger than the crater, with a radius of 2 feet. This pit was filled with water and a 0.5-pound charge was fired just below the surface

of the water. The pit was refilled, and a shot of the same size was fired. The thought underlying the experiment was to apply normal stress to the pit wall, since the water could not transmit shear to this boundary. The transverse motion from the original cratering shot and the first of the pit shots is shown in Figure 11. It is apparent that there has been no marked change in the nature of the motion. The motion of motion from the water shot was obscured by the air blast wave, which reached the instruments before the ground transmitted energy. The Rayleigh wave amplitude from the pit shot increased about 15 per cent in all directions. There is a reduction of *SH* relative to the Rayleigh wave of about 40 per cent, except for instrument 4. The second shot in the water produced an almost perfect reproduction of the wave forms, but the amplitude distribution of *SH* changed markedly, with increases at some instruments and decreases at others. The Rayleigh wave amplitude returned to the same value as it was for the cratering shot.

This experiment has ruled out action within the zone of crushing immediately around



2. Theoretical patterns of radiation of first S motion from propagating rock in an infinite medium (after Knopoff and Gilbert, model 3).

The generating mechanism for SH motion. The normal stresses on the pit walls were uniformly distributed, being zero at the free surface and a maximum at the bottom. The pattern is roughly equivalent to a vertical implosion of the surface, rather than a horizontal expansion. Since the walls of the pit were not significantly altered by the shots, it is assumed that the action takes place in the most highly stressed region of the pit. Unfortunately, it is not practical to carry out an ideal experiment of a spherical pit at great depth. These data indicate that for a large enough depth there would not be a great change in the records already seen for completely vertical shots.

Discussions of SH generation. The discovery of the mechanism by which this energy is released is the ultimate goal of this work. Possible mechanisms are:

Twist about the vertical (or a rotation of the surface of the 'equivalent cavity'). Since the data would be an azimuthally symmetrical pattern, the data on hand rule out this possibility.

Slipping with a horizontal component of motion. Although this is intuitively rejected as a result of an explosion, it has been suggested that such slipping could occur if the shots were fired in a strained medium. However, it is likely that the loess in which these shots were fired can support any great amount of strain over long periods of time. In any

case, the production of almost identical wave amplitudes for repeated shots in the same place, as in the pit experiments, would seem to eliminate the release of accumulated strain as the source. On the other hand, a careful study of the large-scale slipping of blocks of rock as the result of high-yield shots might prove very useful if good azimuthal coverage of ground motion were available. Such slipping has been observed at the Nevada test site, for example, and could be an adequate source of SH motion.

3. Cracking. This may be described more formally as propagating faulting with displacement normal to the fault surface. This is model 3 of Knopoff and Gilbert [1960]. The theoretical radiation pattern for the first S motion from such a source is plotted in Figure 12. This figure is based on equation 12 of the cited paper.¹ The direction of propagation of the fault is to the right. The fault plane extends infinitely and is perpendicular to the plane of the figure, and displacement at the fault surface is up or down on the figure, normal to the fault plane. The plus and minus signs refer to clockwise and counterclockwise motion, and the amplitude scales are arbitrary. The pattern depends on ζ , the velocity of fault propagation—or rate of growth of the crack. To represent SH motion, the cracks would have to be vertical and grow

¹ A factor s_z is omitted from the first term in the expression for U_z . This factor does affect the shape of the radiation pattern.

radially outward. An infinite velocity of cracking yields a symmetric, four-lobed pattern; a velocity twice that of shear waves gives the four-lobed pattern with smaller lobes in the direction of faulting; and for the rate of cracking less than the shear velocity, the two-lobed pattern results.

Radial cracking is a normal result of a contained explosion. Patterns of surface cracking have been sketched wherever they have occurred in the experiments described. Shots at moderate depth that do not remove material to form a crater result in extensive observable cracking. In at least one of the cases discussed here, shot point 24, the most prominent cracks agree well with the nodal lines on the radiation pattern, and, perhaps even more striking, a quadrant free of cracking was found which agrees with the most prominent lobe. In the cratering shots and the shots in the water-filled pits, which behaved much alike, surface cracking was not found, but cracks radiating downward from the crater bottom can be expected.

A preliminary analysis of the data for shots in limestone, in which the total area of crack surface relative to the primary cavity seems to be even larger than in loess, has given additional support to the hypothesis that crack formation is an important source of *SH* energy.

Since cracks radiate in several directions, the final radiation pattern will be a superposition of theoretical curves such as those seen here.

The preferred orientation of the observed patterns indicates that under the cracking hypothesis there is one direction, or more, of weakness, related to the fine structure of the medium, along which cracks form most readily. This condition might ultimately be related to the process by which the soil was deposited.

4. *Energy return from dipping interfaces.* Reflection and refraction of compressional energy and *SV* motion formed by conversion will produce apparent *SH* motion at the surface. The data presented here do not rule out this mechanism. The geology is very uniform over the dimensions of the test area, and regional dips are low. On the other hand, this mechanism would produce a preferred orientation of the radiation patterns, as has been observed at this site. The data indicate that at least the large-

amplitude Love waves are directly produced and do not result from mode conversion.

Conclusions. The data clearly indicate the *SH* motion from small explosions in has an asymmetrical amplitude distribution and reversal of polarity between adjacent lobes.

Long-period *SH* motion (Love waves) is produced when the source is shallow enough relative to charge size, to produce surface craters or cratering. The motion from a complex contained shot is of shorter period and shorter over-all duration than the motion from a shallow shot.

It is suggested that while the classical model of an explosive source as a symmetric center of compression may be adequate to explain compressional waves, and even the Rayleigh waves, it should be recognized as an oversimplification of the complex action around the explosion. The formation of cracks is a normal result of a contained explosion and could yield *SH* motion with the properties observed. The question of how much energy is released into the medium as the result of cracking is a function of the type of material. This is a matter for further study.

Acknowledgments. The research on which this report is based is being done under contract DA-19(604)-7402 with the Advanced Research Projects Agency of the Department of Defense under Project Vela Uniform.

REFERENCES

- Ewing, W. M., W. S. Jardetsky, and F. P. *Elastic Waves in Layered Media*, McGraw-Hill Book Co., New York, 1957.
- Kisslinger, C., Observations of the development of Rayleigh-type waves in the vicinity of small explosions, *J. Geophys. Research*, 64(4), 4436, 1959.
- Knopoff, L., and F. Gilbert, First motion from seismic sources, *Bull. Seism. Soc. Am.*, 50, 117-134, 1960.
- Leet, L. D., Earth motion from the atomic bomb test, *Am. Scientist*, 34, 198-211, 1946.
- Romney, C., Short period shear waves and their application to discriminating between earthquakes and explosions, in *The Need for Fundamental Research in Seismology*, Report of the Panel on Seismic Improvement, U. S. Department of State, appendix 2, 1959.
- Sharpe, J. A., The production of elastic waves from explosion pressures, 1, Theory and empirical field observations, *Geophysics*, 7(2), 144-151, 1942.

Results of the 1960 Expedition to Krakatau

ROBERT W. DECKER

Dartmouth College, Hanover, New Hampshire

DJAJADI HADIKUSUMO

Djawatan Geology, Bandung, Indonesia

Abstract. On January 12-13, 1960, a twelve-man expedition visited Anak Krakatau to record renewed activity. Explosive, vulcanian-type eruptions of pyroclastics from fine ash to blocks 1 meter in diameter occurred at $\frac{1}{2}$ - to 10-minute intervals throughout the period of observation. The repetition of approximately twenty small, turbulent, explosion clouds of gas and ash rising to 300 meters alternated with four to six larger eruptions of gas, ash, and lapilli clouds turbulent rising to 1200 meters, the latter accompanied by larger blocks landing as far as 600 meters from the vent. A new topographic survey was completed, and four maps showing Anak Krakatau's topography since 1950 are presented. Recent fathometer measurements show the total size of Anak Krakatau and suggest that the 1883 caldera floor is being gradually leveled with volcanic detritus. Historic records indicate that the eruptions begin at approximately 200 meters below sea level and continue their way to the surface. Steam generation where the magma reaches the porous, pyroclastic material of Anak Krakatau is considered an important contribution to the periodic gas explosions. The energy released in individual large explosive eruptions is estimated to be 170 tons of TNT equivalent. Over 20-minute increments, the rate of energy release is nearly constant at 3.1×10^{13} joules, or 31,000 tons of TNT equivalent, per day. The present composition of the ejecta is still highly silicic, and, although no danger of further collapse is indicated, more systematic observation of renewed activity is warranted for scientific purposes.

Introduction. In June 1959, Anak Krakatau, a growing volcanic cinder cone on the rim of the 1883 Krakatau caldera, again showed renewed activity [Decker, 1959]. This was the first substantial activity since eruptions during 1883 [Neve, 1956]. After several weeks of eruptions the activity ceased, and it did not begin again until December 1959, when more eruptions renewed the growth of Anak Krakatau. Mindful of the tidal waves from the 1883 caldera collapse which killed 36,000 people in 1883, the Indonesian Government asked the Indonesian Geological Survey to send a Geology Department at the Institute of Geology Bandung to ascertain the possibility of renewed explosions or caldera collapse of the same magnitude, and also to continue the record of eruptions which have intermittently been observed at Anak Krakatau since its submarine eruptions in 1927.

A twelve-man scientific party visited the caldera on January 12 and 13, 1960, in a 20-ton police boat, which provided not only reliable transportation but a convenient platform for observations.

of eruptions. The entire cone of Anak

Krakatau has been built of pyroclastic material, and the January 1960 eruptions were no exception. Explosive eruptions occurred at 30-sec to 10-min intervals throughout the 2 days of observation. The eruptions are perhaps best described as vulcanian in character, but the regular repetition of the explosions and lack of bread-crust bombs do not fit the classical description of an eruption of vulcanian type. However, the dense, rapidly rising, turbulent explosion clouds are similar to vulcanian eruptions, and the high gas content and gas pressures needed to form these explosion clouds are certainly present.

The pyroclastic ejecta range from fine ash to 2-meter angular blocks of porphyritic basalt, the blocks being thrown beyond the outer crater rim to a maximum of 600 m horizontally from the vent during the larger eruptions and the fine ash being occasionally carried downwind as far as the coast of Java, 50 km distant. Individual explosions showed a great range of violence from small gas jets less than 30 m high to large, turbulent mushroom clouds over 1000 m high. The noise of the explosions, even in the largest eruptions, was not impressive, sounding much



Fig. 1. 300-meter eruption viewed southward from the outer crater rim. The rim of the active cinder cone is 100 meters in diameter.

like a deep and effective quarry blast at the visual beginning of the explosion and then becoming almost noiseless as the turbulent explosion clouds grew upward. The maximum height of the *turbulent* clouds gave the best index of the size of the individual explosions. The height at which turbulence stopped is stressed because, depending on wind conditions, the eruption clouds drifted much higher or were dissipated after the turbulence ceased but were little affected by winds during their turbulent climb. Since turbulence partially ceases when the gas inside the explosion cloud reaches pressure and equilibrium with the atmosphere, the maximum turbulent height also makes possible a rough estimate of the gas volume in the individual explosion clouds.

A cyclic repetition of approximately 20 small explosion clouds, whose turbulent heights would seldom exceed 300 m (Fig. 1), with almost no blocks ejected, alternated with 4 to 6 larger explosions (Fig. 2). The cyclic relationships are

best seen in Figure 3. In this graph each vertical line represents a single explosive eruption; the top of the line indicates the maximum height of the turbulent cloud associated with the eruption. The horizontal scale records the time of each explosive eruption. An arrowhead at the top of a vertical line indicates that a show block was thrown beyond the outer crater rim. Periods of 2 to 10 min separate the larger eruptions, and periods of 30 sec to 2 min separate the smaller episodes. The alternations of small and large eruption cycles were evident in early observation on January 12, but the regularity of this cycle became remarkable during most of January 13. The sinusoidal character of the increasing and decreasing violence of *individual* eruptions, shown by the graph, was striking. The period ranged from 50 min to 1 hour from peak to peak of maximum explosive activity. This cycle has no apparent diurnal or tidal relations and is considered to be an inherent cycle of the energy flux and geometry of the present Krakatau volcanic system.

Although the individual explosive eruptions were of quite different magnitude, the energy released by Anak Krakatau over intervals of time covering several eruptions was almost constant. In effect, a longer delay between eruptions meant that more pent-up energy would be released when the next eruption threshold was reached. This was not precisely the case for individual eruptions because a delay of 4 min would sometimes be followed by a large eruption as followed a delay of 10 min. However, if the estimated energy release in individual eruptions over any period of 20 min is summarized, the result is nearly constant (Fig. 4).

Topography. Anak Krakatau, as surveyed in 1960, is a semicircular island with a minimum NE-SW diameter of 1.5 km. Its high point on January 12, 1960 was 166.7 m on the outer crater rim. The outer crater rim associated with the old crater lake vent is 600 m in diameter and the recent inner cinder cone is 100 m in diameter and 50 m high.

The entire cone is asymmetrical to the southwest, probably owing in part to the large submarine volume which has been filled in on the southwest side. The vent of Anak Krakatau is on the northeast rim of the 1883 caldera; the deep water of the caldera bottom was on the southwest side of the vent. The northeast



600-meter eruption viewed from the west shore of Anak Krakatau Island.

tively shallow platform; therefore more construction on the northeast has been. This is augmented by the strong wind and surf from the Indian Ocean which erode the loose unconsolidated pyroclastic material from the south and southwest sides of the island. These same currents carry the material to the north shore of the island where small spits are beginning to form on the northeast and southwest corners of Anak Krakatau.

Figures 4, 5, 6, and 7 show the changes in the crater lake which have taken place on Anak Krakatau since 1950. The record of activity and the crater lake from the beginning in 1927 to 1941 is given [Neuman van Padang, 1951], but from 1941 to 1950 the record is sketchy owing to the political upheavals in Indonesia. Since 1949 have been made by the Indonesian Geological Survey. The 1950 map shows that during a considerable period of little volcanic activity erosion encroached well into the crater lake, completely breaching the southwest rim. The present crater lake is separated from the

sea by only a very low sand bar. Renewed activity in 1952 and 1953 [Neve, 1956] rebuilt the southwest crater rim, as is shown in the 1952 and 1953 maps, but all the eruption products of this activity, as well as earlier eruptions, came from beneath the crater lake. An interesting sidelight on the 1950, 1952, and 1953 maps is the height of the crater lake. The lake level is shown to be above sea level by 2.79, 3.4, and 0.94 m, respectively. Anak Krakatau is composed entirely of pyroclastic material, much of it sand-size, which is quite obviously a porous and permeable medium. The lakes must therefore be in hydraulic communication with the sea. A rather large Ghijben-Herzberg lens of fresh or brackish water must exist in the water-table system of Anak Krakatau, floating on more dense sea water beneath. The vegetation at sea level on the north spits of the island substantiates this conclusion, which certainly is not unusual in an area of more than 200 cm of rain per year. The interesting point is not the fact that the lens exists but the manner in which the lake surface level discloses the phenomenon.

The main change shown on the 1960 map is the development of the new inner cinder cone which formed during the 1959-1960 activity. This cone has nearly displaced the old crater lake, and it gives Anak Krakatau a definite subaerial vent for the first time. Whether this vent is smaller owing to less vigorous activity than during the crater lake eruptions is questionable. The larger diameter (600- to 700-m crater rim) of the crater lake stage may have been due to the erosional power of the agitated lake surface rather than to greater energy of eruptions than at present. It would be interesting to know whether other sub-lake vents, in comparison with subaerial vents, show this same sudden change in diameter of their associated craters.

Fathometer traverses were made with the echo sounder on the police boat over the 1883 caldera and around Anak Krakatau. These were made to compare the present bottom configuration with earlier measurements and to aid in determining the present volume of Anak Krakatau. This new volume, representing material ejected since 1927 when Anak Krakatau began its submarine eruptions, is computed to be 0.3 km³. The present bulk of Anak Krakatau represents approximately 5 per cent of the missing, collapsed volume of the pre-1883 Krakatau.

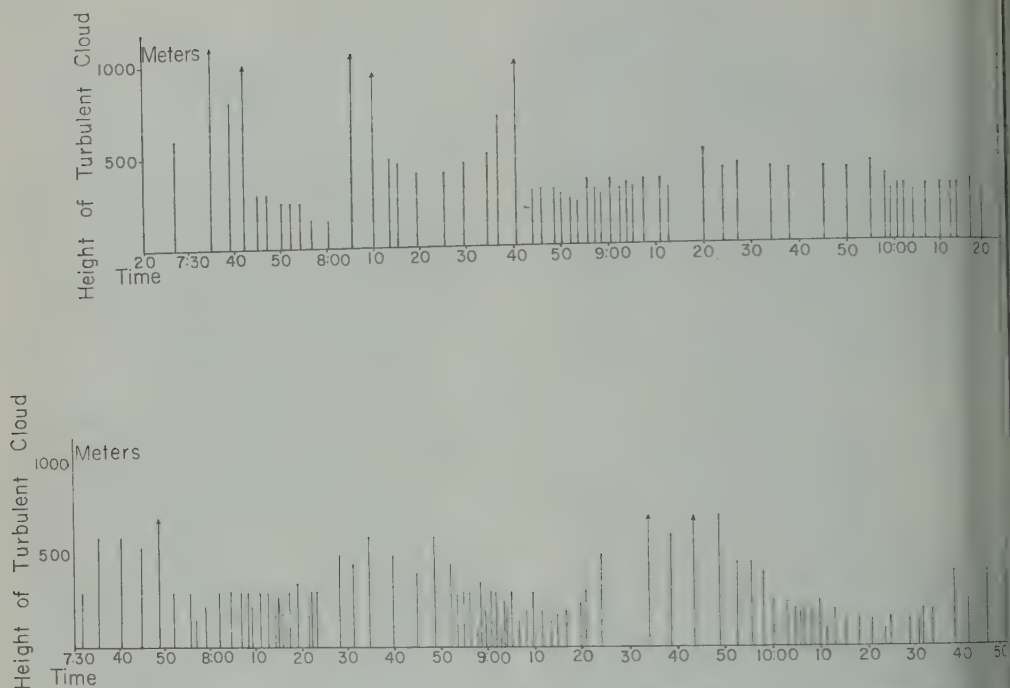


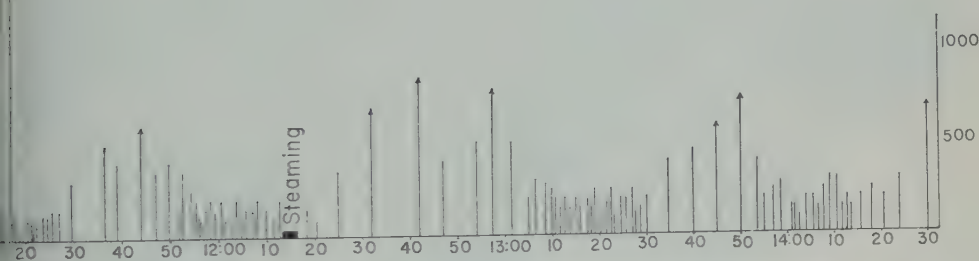
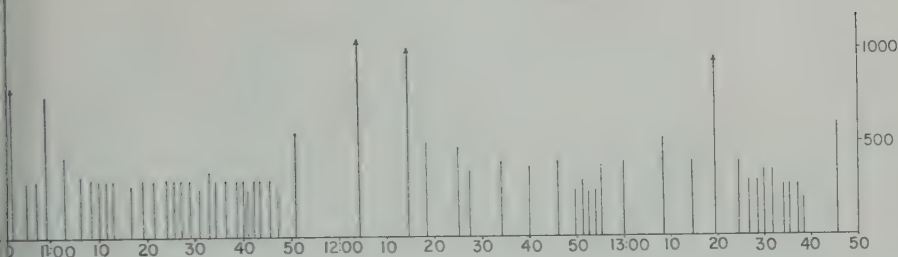
Fig. 3. Activity of Anak Krakatau (a) January 12, 1960, (b) January 13, 1960. Hours given in (a) are in Indonesian time.

Continued growth at the rate of the past 33 years would require at least 600 more years for Anak Krakatau to replace the volume lost in the 1883 catastrophe, a rapid rate indeed relative to most geologic phenomena.

Rather drastic changes took place in the bottom configuration of the 1883 caldera between early soundings made shortly after 1883 and more systematic soundings made in 1923 [Neuman van Padang, 1933]. Increases in depth were noted from a few meters to a maximum of 196 meters. Williams [1941, p. 264] believed that such apparent vertical sagging may have been caused by volume change of residual magma crystallization, by magma movement toward the incipient vent of Anak Krakatau, or by settling of the collapsed blocks in the caldera. The recent fathometer traverses across the deeper parts of the 1883 caldera indicate that the subsidence has stopped. On a line from the high point on Sertung Island to Zwarte Hoek on Rakata Island the deepest point was 238 m (Fig. 8). The depths in this area exceeded 250 m in the 1923 survey. Rather than being attributable to vertical uplift, these post-1923 changes can readily be explained as sediment filling derived from water-

lain ash and submarine slides from Anak Krakatau. The very flat bottom of the present caldera floor favors the depositional explanation.

Seismic data. Records of Spindler and Hopy, vertical and horizontal mechanical seismographs placed 995 meters from the active vent on the north shore of Anak Krakatau, were obtained on January 12 and 13, 1960. Several damping ratios were used, and magnifications were from 67 to 129. Surface waves were detected with the smallest explosive eruptions, and distinct first arrivals of compressional waves were detected from most of the larger and intermediate explosions. The surface waves obscured useful data after the first arrivals, but interesting results can be interpreted from the first arrivals alone. The horizontal component was aligned approximately radially to the explosion vent and results were corrected to a true radial position. The ratios of corrected amplitudes of the vertical to the horizontal component for the first arrivals ranged from 0.28 to 0.31, giving an average apparent angle of 74° between the radial and the vertical. Combining Byerly's [1942, 1943, 1948] formula for reflection of incident rays at a free surface with Dobrin's [1960, p. 20] formula



Djakarta) time. Arrows indicate that blocks were thrown beyond the outer crater rim.

compressional and shear waves to ratio yields

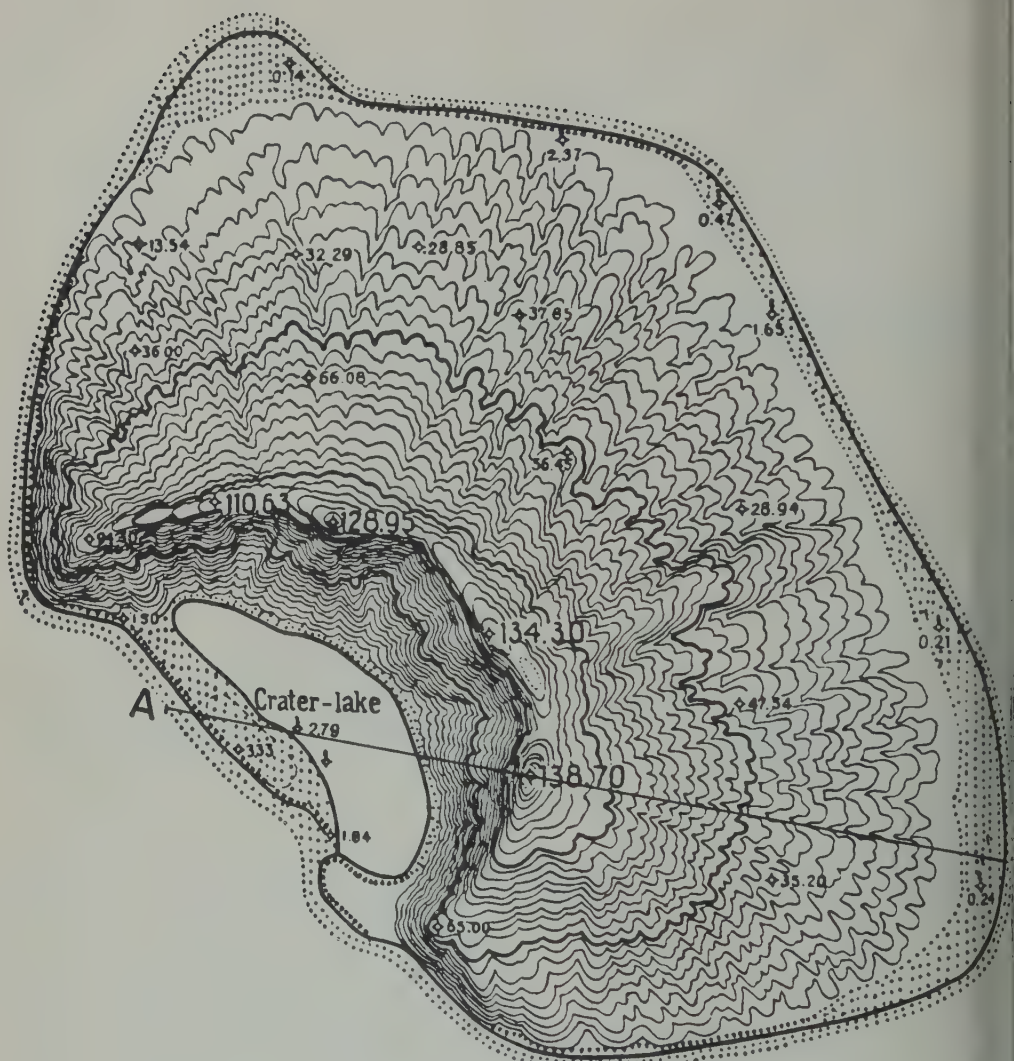
$$\alpha = \left(\frac{1 - \sigma}{\frac{1}{2} - \sigma} \right)^{1/2} \left(\frac{1 - \cos \bar{\alpha}}{2} \right)^{1/2}$$

is the apparent angle of incidence, α is angle of incidence, and σ is Poisson's ratio of the incident medium. Values of Poisson's ratio from Birch's [1942, pp. 73-77] handbook for approximating the unconsolidated rocks of Anak Krakatau range from 0.11 to 0.2, with a mean value of 0.2. For a Poisson's ratio of 0.2, the computed true angle of incidence is probably not a valid assumption because of density changes and layering, the incident ray intersects a vertical line through the active vent at 205 m below sea level, the apparent point of origin of the subsurface explosions. Owing to the high degree of uncertainty of Poisson's ratio and to the fact of the convex upward refraction of the ray probably taking place, this depth is at best a rough approximation. Even if the actual depth of the explosions is twice the

calculated depth, i.e. 400 m, this is still a very shallow depth for volcanic activity; it suggests a surficial mechanism for the explosive phenomena. When the seismic cross section (Fig. 9) is compared with the same geological cross section (Fig. 10), there is good reason to believe that the 200-m-depth interpretation is not in great error, since this is just below the intersection of the 1883 caldera rim fault with the base of Anak Krakatau, a point where phreatic explosions could be generated.

The interval between first arrivals and the initial surface explosions ranges from 15 to 25 sec, even though the amplitude ratios of the vertical versus horizontal seismic components indicate nearly identical points of origin of the subsurface explosions. This lag is inferred to be the time it takes the subsurface explosion to churn its way to the surface, and the variation in this lag may be due to random settling and packing following each explosive episode.

Mechanism of explosive eruptions. Unfortunately no equipment for gas analysis was available. The explosion gases smelled only slightly of sulphur dioxide and were not toxic during brief encounters with downwind remnants



H and V Scale

0 300 600

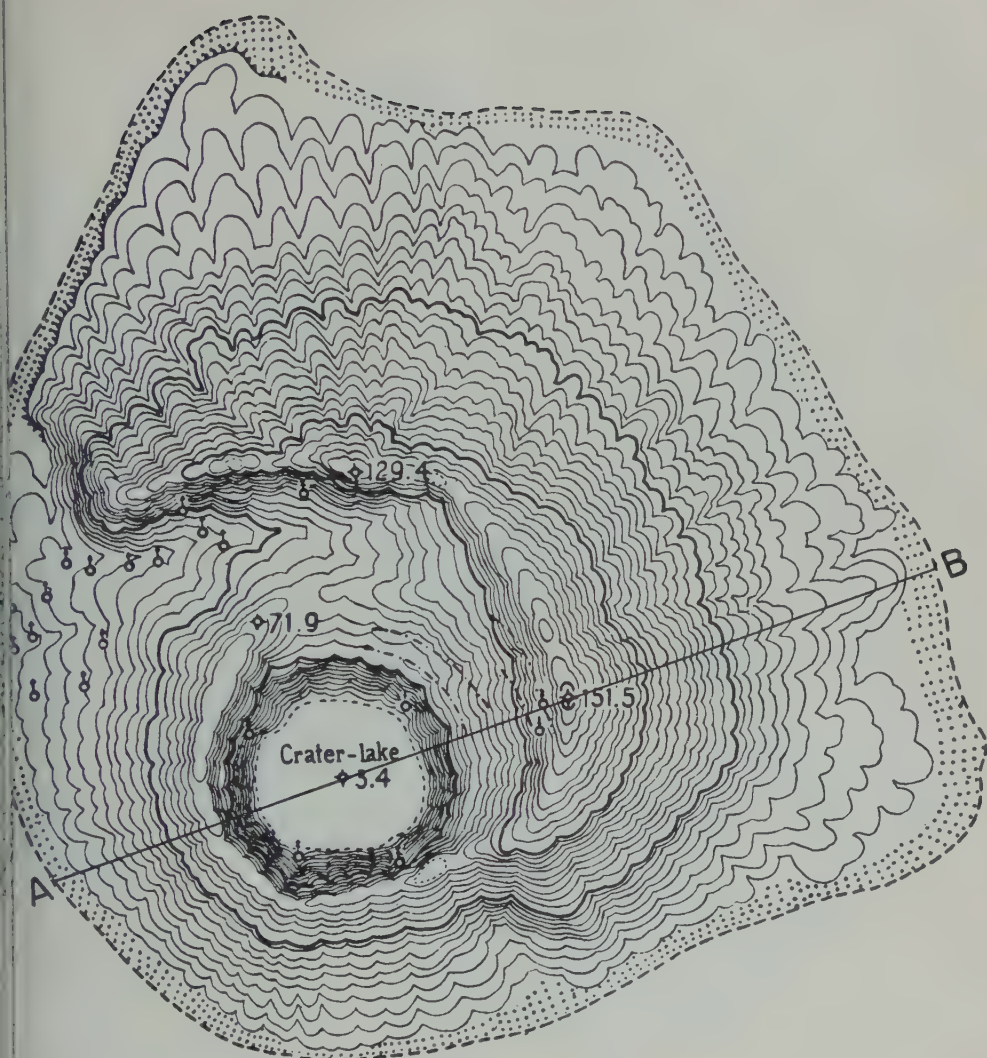
Meters

2.79
Crater lake

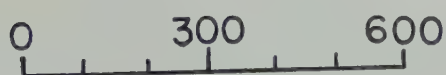
138.70

A

Fig. 4. The state of Anak Krakatau in 1950 (surveyed by the Topographical Survey in September with corrections of the Volcanological Survey).



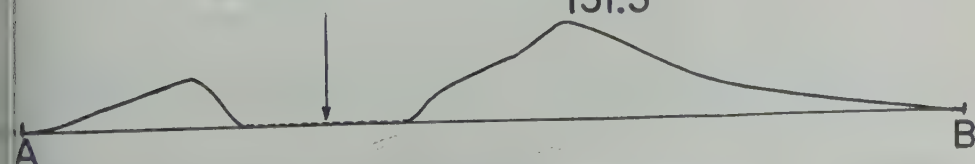
H and V Scale



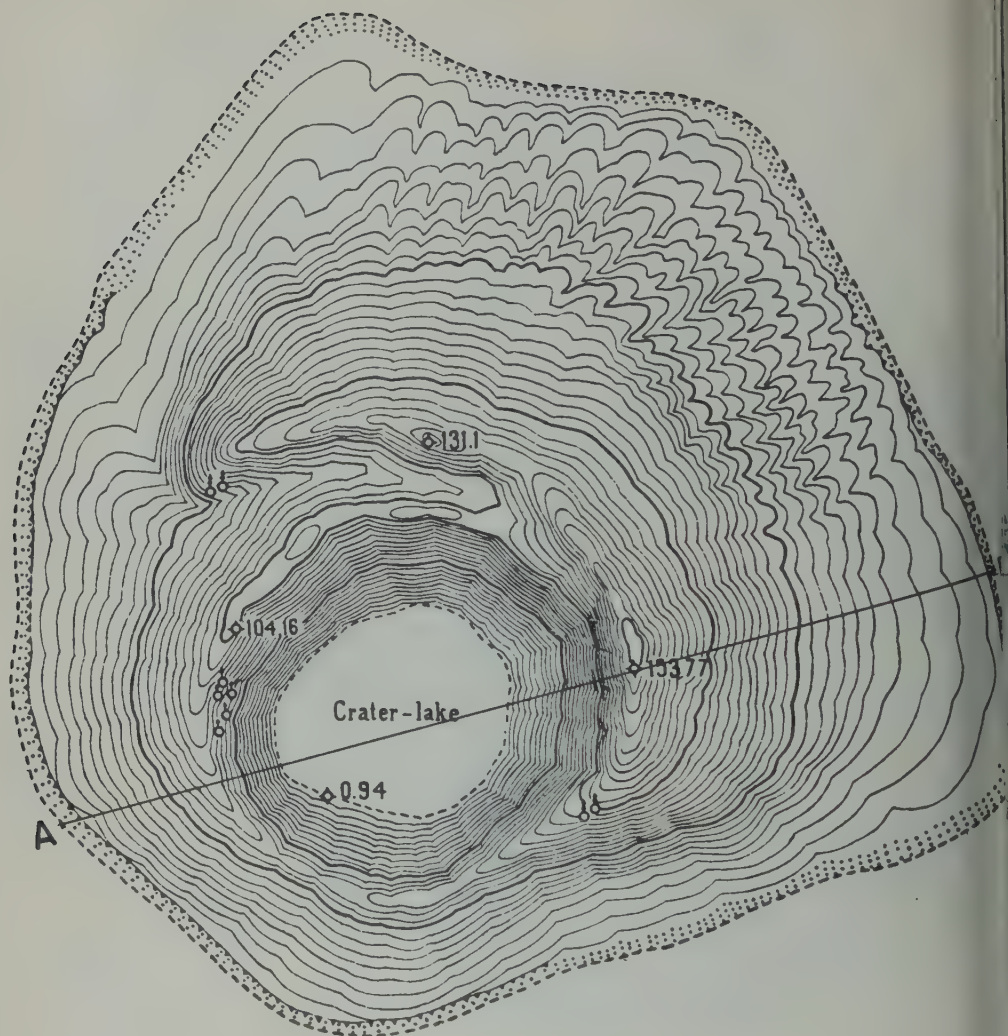
3.4 Meters

Crater lake

151.5



5. The state of Anak Krakatau in 1952 (surveyed by the Volcanological Survey in October).



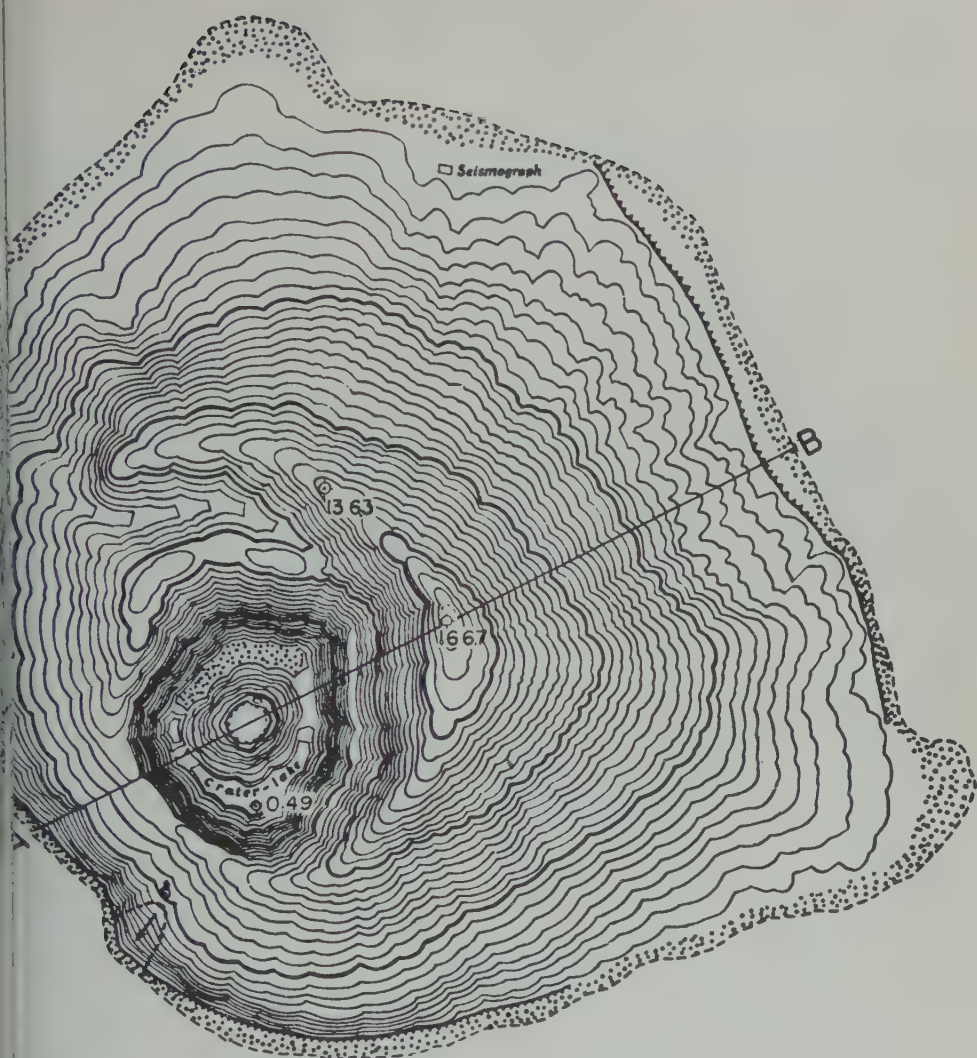
H and V Scale
 0 300 600
 Meters

0.94
 Crater lake

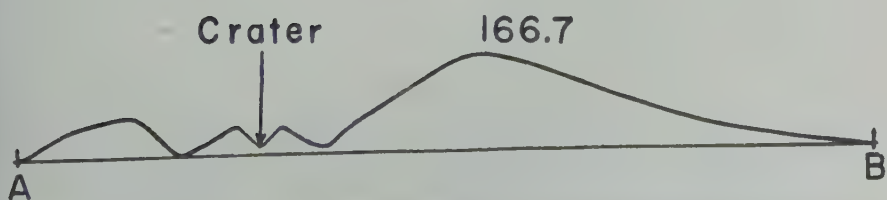
153.77

A

Fig. 6. The state of Anak Krakatau in 1953 (surveyed by the Volcanological Survey in October).



H and V Scale
 0 300 600
 Meters



The state of Anak Krakatau in 1960 (surveyed by the Volcanological Survey in January).

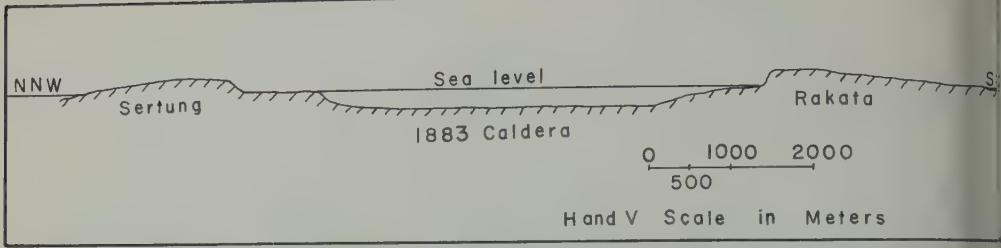


Fig. 8. 1960 profile of the 1883 caldera floor from fathometer soundings between high point on Sertung Island and Zwarte Hoek on Rakata Island.

of the explosion clouds. From these admittedly thin data it is inferred that the explosion gas is largely steam in even greater proportion than the normally high steam content of volcanic gas. This abundance of steam is difficult to reconcile with basaltic magma. Although no experimental data similar to *Goranson's* [1931] work on solubility of water in granitic magma are available, most petrologists suppose that basaltic magma is less hydrous than granitic [*Foshag*, 1950, p. 752], and *Bowen* [1928] argued for comparatively anhydrous basaltic magma. *Foshag* estimated that over 17 per cent of the emitted products during eruptions of Paricutin was water vapor, and he suggested a considerable dilution of magma emission by vapors derived from meteoric waters.

The basaltic blocks thrown out during the

recent activity of Anak Krakatau show less than 5 per cent porosity of random vesicles, which range from $\frac{1}{2}$ to 5 mm in size. The cooling of these blocks must have been at a sufficient shallow depth to allow these few vesicles to form, yet their small volume suggests that the magma itself is not rich in dissolved gases or water vapor. The alternative conclusion that the steam is largely of nonjuvenile origin overcomes these difficulties and is in agreement with the geophysical and geologic evidence.

The mechanism of the recent explosive eruptions of Anak Krakatau is hypothesized as follows: New magma is entering the remains of the collapsed 1883 reservoir. Some leakage of this basaltic, low-vapor-content magma is escaping up the rim fault of the 1883 caldera. This escape is taking place in small, 30-sec-

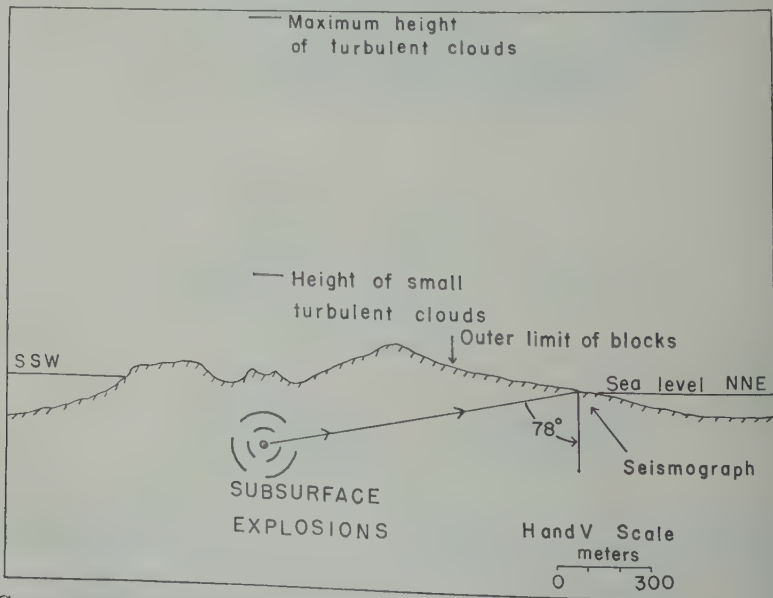


Fig. 9. Cross section of Anak Krakatau showing seismic interpretation of explosive eruptions.



Fig. 10. Geological cross section of Anak Krakatau.

sations owing to build-up and release of low vapor pressures. The intervals of pulsations also vary in nearly hourly cycles (especially on January 13, 1960) because of an unknown mechanism of restriction in the conduit, or because of the over-all geometry of the input of the system. With each pulse of liquid magma into the zone of sea level at or below the base of the impermeable bulk of Anak Krakatau, generation of phreatic steam gives rise to surface gas explosion. The explosion is directed away to the surface and blows out in a conical cinder cone. It carries upward ash and fragments of the suddenly chilled basaltic magma, and some blocks of the more slowly moving magma of the 'dry' conduit lining, which have been ripped loose and churned up by a series of the explosions. It should be noted that other hypotheses of mechanism are not supported by the meager available data. The above model merely represents the authors' present interpretation and serves as a working model to facilitate further investigation and stimulate discussion.

A glassy skin which forms on basaltic magma when it is exposed to the open sea greatly retards rapid heat transfer and this forms a formidable objection to the generation of steam necessary for the phreatic explosions. However, the glassy skin melted out in the recent activity of Anak Krakatau. The ash is porphyritic, and the ash contains plagioclase crystal fragments, indicating that magma is leaking up the caldera rim. The ash is already a crystal mush. Under these conditions, the glassy skin caused by sudden cooling would probably give way to intense heat transfer and the sudden, explosive

generation of steam. Palagonite is present in the newly erupted ash, but the blocks show no development of palagonite. This is good petrographic evidence that water contamination has taken place, but it does not establish whether the time of absorption was before or during the explosions.

If steam generated largely from surface waters caused the recent explosive activity of Anak Krakatau, the volcano belongs in *Williams'* [1954, p. 322] phreatomagmatic classification. This does not imply, however, that a phreatomagmatic mechanism explains the catastrophic eruptions of Krakatau in 1883. The great volumes of pumice involved in that eruption suggest a most intimate solution of water within the magma before pressure release allowed the rapid vesiculation and expansion of the glassy froth. The careful work and conclusions of *Stehn* [1929], based on his own investigations as well as on the original great Krakatau monograph of *Verbeek* [1886], are in no way altered by the present interpretations of eruption at Anak Krakatau.

Energy of eruptions. Energy is released during volcanic eruptions in so many forms that attempts to compute the total energy output are only approximations at best. Nevertheless, such attempts are worth while in clarifying thinking on the manner and importance of the diverse ways in which energy escapes from a volcanic system. Several of the possible energy release mechanisms are considered below, with respect to their importance in the recent activity of Anak Krakatau.

Increase in local thermal gradient and heat flux: Soil temperatures and the sea water temperature around Anak Krakatau (26° to 29°C) are not abnormally high for this region

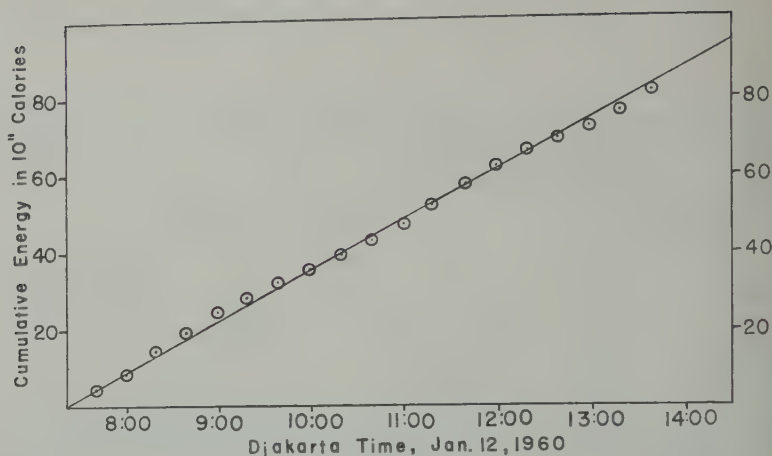


Fig. 11. Estimated cumulative energy release from Anak Krakatau on January 12, 1960. The data points represent energy summations of individual explosive eruptions into 20-minute intervals and show a nearly constant rate (straight line) of energy release. Extrapolation of this graph indicates an energy release of 3.1×10^{13} cal/day.

and suggest that very little heat loss over and above the background heat flux is taking place in this manner.

Seismic energy: Several computations relating the ground motion, explosive energy, and seismic magnitude and energy can be made using empirical formulas derived from underground explosion tests. The large explosive eruption of Anak Krakatau at 12:14 Djakarta time, January 12, 1960, is selected for such analysis. Amplitude of the maximum measured ground displacement was 240μ , and the distance from the seismograph to the explosion center was approximately 3400 feet. *Carder and Cloud* [1959, p. 1474] established the relationship between ground motion amplitudes and the total energy of explosions as

$$A = 3.6(W^{0.75}/D^2) \times 10^4$$

where A is the zero to peak ground displacement in centimeters, D is the distance in feet from explosion to recorder, and W is the explosive energy in tons of TNT equivalent. According to this formula, the amount of TNT detonated to produce similar ground amplitudes would be 15.3 tons, or 1.53×10^{10} cal. *Romney* [1959, p. 1498] derived the empirical formula

$$M = 3.65 + \log Y$$

from seismic recording of several underground nuclear explosions, where M is the magnitude of an earthquake which would produce equivalent seismic waves and Y is the energy of the explo-

sion expressed in kilotons of TNT equivalent. The earthquake magnitude produced by 12:14 explosion of Anak Krakatau should therefore be 1.83. Unfortunately, the Djakarta seismograph station was not able to confirm the occurrence of seismic waves from this explosion of Anak Krakatau and thus giving no independent magnitude computation. Seismic energy of an earthquake can be calculated using *Gutenberg and Richter's* [1956] formula

$$\log E = 9.4 + 2.14M - 0.054M^2$$

where E is the energy in ergs and M is the earthquake magnitude. The seismic energy of 12:14 Anak Krakatau explosion was then 1.38×10^{14} ergs or 3.3×10^6 cal, or only 2% of the explosive energy.

Generation of steam: Because rapid steam generation is considered an important mechanism in the recent activity of Anak Krakatau, considerable energy must be expended to generate this gas. To calculate this energy, the volume of steam generated in each explosive eruption must first be estimated. The clouds grow rapidly until the steam has expanded and cooled to the condensation point. Total cloud volume is roughly the cloud volume at the time turbulence ceases, and it can be approximated by the cylindrical shape of the eruption clouds. The cloud diameters average nearly 100 m, and their heights are a function of the explosion intensities, ranging up to 1100 m in total to

Therefore terminal cloud height is a function of eruption energy as far as steam is concerned, and this relationship is used in the computations for Figure 11. For the 12:14 eruption of January 12, 1960, assuming the shape of this cloud to be approximately a cylinder 100 m wide and 100 m high, containing a volume ($\pi r^2 h$) of 3.14×10^6 m³. A cubic meter of steam contains 0.8 g of water at atmospheric pressure (1 atm). Therefore, the maximum water content of such a cloud would be 4.7×10^6 kg. It is assumed that all the steam in the cloud condenses. This is a continuous process during eruption cloud growth; air is pulled into the condensing steam, and it maintains its buoyancy. Most of the condensing water is immediately absorbed in the displacing air, and the heat is carried away by the caloric exchange. The heat necessary to vaporize 4.7×10^6 kg of water from 30°C is 609 cal/g, or an amount of 2.86×10^{12} cal. Obviously, some heat is not required to displace the steam in the eruption clouds takes place, but gas analyses from the clouds this is uncertain. However, the error caused by allowing for mixing will make the energy figure too large. On the other hand, the sub-steam is heated well above 100°C, and requires additional energy input. Ejected blocks at temperatures of 300°C. To form steam of this temperature would require an additional 16 cal/g energy input. For the 2.86×10^{12} caloric figure to be in error of magnitude, the amount of magma required to form this steam on contact with water can also be estimated. The nature of the basalt blocks and the fragments in the ash and lapilli indicate magma injected beneath Anak Krakatau is half crystallized, probably at a temperature of 1150°C. The average heat capacity is 0.3 cal/g over the cooling range from 1150°C to 300°C, and the computed heat of crystallization is approximately 86.2 cal/g [Birch, 1942, p. 18]. For the necessary steam assumed for the 12:14 explosion cloud, 9.6×10^6 kg of crystallized magma quenched from 1150°C to 300°C would be required. This figure is quite high. It would take only 200 m³ of pyroclastics of this volume to form a cinder cone—less than 1 day at

the recent frequency of eruption. To build the entire 0.3-km³ bulk of Anak Krakatau would require some 80,000 eruptions of this size—about 250 days of eruptive activity of the intensity of January 12, 1960. Both of these growth rates are too high by a factor of about 20, judging from the eruptive history of Anak Krakatau.

Either the initial estimate of the steam volume in the explosion cloud is too high or a very large portion of the steam must be of magmatic origin. The first alternative would require the unknown air mixing factor to dilute the cloud to over 95 per cent air, and this seems to be the most desirable hypothesis. A reduction of the meteoric water volume needed to produce the steam by a factor of 20 for the air dilution in the cloud and an additional 2.6 per cent reduction for the heat used to raise saturated steam from 100°C to 300°C gives a revised estimate of 2.3×10^6 kg of water converted to steam in the 12:14 eruption. The revised total energy of steam generation would be 1.44×10^{11} cal, 1.40×10^{11} cal necessary for vaporization and 0.04×10^{11} cal in the additional heating above 100°C. The total heat would be supplied by cooling 4.8×10^6 kg of magma, an amount in reasonable accord with the growth rate of Anak Krakatau.

Mechanical work: The explosive drilling of the eruptions to the surface, the energy in the ejected blocks, and the turbulent lift of the ash cloud into the air are largely the result of the energy released by the expansion of the high-temperature steam. Heat equivalent of this potential external work is 10^{10} cal, the original steam temperature being approximately 300°C on the basis of temperature of ejected blocks. The maximum horizontal range of the blocks thrown by the larger eruptions was 600 m. Since the horizontal range is maximum at 45°, neglecting air resistance, $R = V^2/g$ at maximum range, where R is the range, V is the initial velocity of the ejected block, and g is the acceleration due to gravity. Blocks thrown this distance would require initial velocities of 77 m/sec. Comparing this with Minakami's value of bomb velocities of 212 m/sec associated with gas pressure of 563 atm at Asama volcano [Williams, 1954, p. 319] and scaling it down by the interior ballistics formula derived by Hayes [1938, p. 77] yields

$$P'v = 4wa^2/27gAb$$

where $P'v$ is the maximum pressure-producing velocity only, w is the weight of the projectile, a is the theoretical value of the velocity of the projectile attained in a gun of infinite length, g is the acceleration due to gravity, A is the area of cross section of the gun bore, and b is a constant of the particular system. Since w , g , A , and b can be numerically handled as a new constant of 0.0846 in relating Minakami's pressure and velocity values, the equation reduces to

$$P'v = 0.0125a^2$$

for $P'v$ in atm and a in m/sec. For the 77 m/sec maximum-velocity blocks at Anak Krakatau, the corresponding gas pressure would be 74 atm. Saturated steam at this pressure would have a temperature of 292°C, which is in good agreement with the temperature of the ejected blocks. The energy required to hurl a 100-kg block the maximum range of 600 m would be $mv^2/2$, or 7.08×10^4 cal. All the 4.8×10^6 kg of pyroclastics in the revised estimate of the 12:14 eruption could be raised from 200 m below sea level and hurled to this maximum range with only a fraction of the estimated potential energy in the high-pressure steam. However, great frictional losses and energy consumed in rotating and churning subsurface blocks must be overcome as the explosion drills its way to the surface, so that only a small fraction of energy would be finally available to produce the spectacular showers of blocks and the turbulent mushroom explosion clouds.

Final cooling of the pyroclastics: The remaining source of energy release is the cooling of the ejected pyroclastics to atmospheric temperatures, that is from 300° to 30°C. The heat capacity of basaltic material over this temperature range is 0.25 cal/g [Birch, 1942, p. 235], so the energy release would be 3.24×10^{10} cal for the pyroclastics of the 12:14 eruption of January 13, 1960.

Table 1 is a summary of the estimated energy released in a typical larger eruption of Anak Krakatau during the recent activity. The total estimate of 1.7×10^{11} cal is in fair agreement with the explosive energy of 1.53×10^{10} cal computed from the seismic data. The steam-generation mechanism would be less rapid than the energy release from detonating TNT, and

the associated explosion and seismic count should be less efficient. The bomb and perhaps gives a more meaningful quantity in calories, and the total estimated energy release from Table 1 (1.7×10^{11} cal) would be equivalent of an explosion of 170 tons of TNT about the size of a very small nuclear bomb.

Figure 11 summarizes the cumulative energy released by Anak Krakatau over a 7-hour period on January 12, 1960, in increments of 20 minutes. The straight-line nature of the graph confirms the impression that the energy release rate is nearly constant if considered over an interval of several individual eruptions. Extrapolation of this graph gives a daily energy release of 3.8×10^{13} cal or 31,000 tons of TNT equivalent!

Petrography. Neumann van Padang [1932, pp. 62–63] stated that the eruption products of Anak Krakatau were basaltic during 1927–1935 and andesitic since 1935. The recent eruption products do not substantiate this alleged change. Analyses of the 1960 eruption products show silica contents from 49.4 to 53.2 per cent, which is considerably less than the silica content of average andesite [60.3 per cent, according to Daly, 1933, p. 447]. Nevertheless the analyses are somewhat anomalous to the apparently normal basaltic composition of a 1960 block identified in thin section. The section consists of 30 per cent plagioclase phenocrysts (An₁₀), 10 per cent augite, accessory olivine, and a groundmass of microcrystalline plagioclase-augite and glass. Many of the plagioclase phenocrysts show inclusions of glass and lithic debris, suggesting considerable contamination of the basaltic magma by earlier eruption products of Krakatau. Perhaps these inclusions cause the slightly higher and variable silica content of the rock analyses.

The recent ash consists of approximately 60 per cent crystal fragments and 40 per cent fragments of microcrystalline matrix and glass. Chemical analysis of this ash yields 49.4 per cent silica; yet the light green to light yellow-brown color of the glassy fragments, and the low refractive index of 1.54, are not typical of unaltered basaltic glass. Instead they indicate palagonitization. Chemical analysis of the ash also shows combined water content of +2.1 per cent, apparently related to the palagonitization since combined water is less than 1 per cent in the volcanic blocks.

The recorded geologic history of Krakatau

Estimated Mechanism and Amount of
Released by the 12:14 Explosive Eruption
Anak Krakatau, January 12, 1960

Energy Release	Energy, cal	Percentage of Total
of high temperature steam from boiling water	1.4×10^{11}	82
rejected pyro-	0.3×10^{11} 1.7×10^{11}	18

close relationship between two periods
the collapse that are associated with
magma, and renewed but non-
eruptive activity of basaltic and
residual materials. Composition change
more siliceous magma is therefore a
signal of approach to a possible third
catastrophic caldera collapse. Since the
products of Anak Krakatau are still
there seems to be little danger of
caldera collapse within the next few
years. At the growth rate of Anak
since 1927, another 600 years will be
to rebuild the missing volume of the
crater, and, if future activity is con-
tinued with past compositional changes, large
amounts of andesite and finally small warning
of siliceous pumice should precede
collapse.

Although a future disastrous eruption seems
the present activity of Anak Krakatau
intense scientific observation. If con-
servative observation is not possible, periodic
expeditions to the island are certainly
needed. Anak Krakatau is a potential
geochemical, and geophysical labo-
ratory without comparison. The well-known
geologic history of the volcano puts
it into an excellent frame of reference,
the system seems to be so dynamic that the
chemical chemistry of the eruptive mechanism
must be understood.

REFERENCES

- Birch, Francis (Ed.), Handbook of physical con-
stants, *Geol. Soc. Am. Spec. Paper 36*, 325 pp.,
1942.
- Bowen, N. L., *The Evolution of the Igneous Rocks*,
Princeton University Press, 334 pp., 1928.
- Byerly, Perry, *Seismology*, Prentice Hall, New York,
256 pp., 1942.
- Carder, D. S., and W. K. Cloud, Surface motion
from large underground explosions, *J. Geophys.*
Research, 64, 1471-1487, 1959.
- Daly, R., *Igneous Rocks and the Depths of the Earth*,
McGraw-Hill Book Co., New York, 598 pp., 1933.
- Decker, R. W., Renewed activity of Anak Krakatau,
Inst. Technol. Bandung, Contrib. Dept. Geol., 34,
pp. 1-5, 1959.
- Dobrin, M. B., *Introduction to Geophysical Pros-
pecting*, 2nd ed., McGraw-Hill Book Co., New
York, 446 pp., 1960.
- Foshag, W. F., The aqueous emanation from
Paricutin volcano, *Am. Mineralogist*, 35(9, 10),
749-755, 1950.
- Goranson, R. W., The solubility of water in granite
magmas, *Am. J. Sci.*, 22, 481-502, 1931.
- Gutenberg, B., and C. F. Richter, Earthquake
magnitude, intensity, energy and acceleration,
Bull. Seism. Soc. Am., 46, 105-145, 1956.
- Hayes, T. J., *Elements of Ordnance*, John Wiley &
Sons, New York, 715 pp., 1938.
- Neumann van Padang, M., De Krakatau voorheen
en Thans, *De Tropische Natuur.*, Jaarg. 22,
Afl. 8, 137-150, 1933.
- Neumann van Padang, M., *Catalogue of Active
Volcanoes of the World Including Solfatara Fields*,
1, *Indonesia*, International Volcanological Assoc.,
Napoli, Italy, 1951.
- Neve, G. A. de, Krakatau and Anak Krakatau with
a communication on the latest investigation in
October, 1953, *Proc. Eighth Pac. Sci. Congr.*,
1953, 2, 178-179 (abstract only), Quezon City,
Philippines, 1956.
- Romney, C., Amplitudes of seismic body waves
from underground nuclear explosions, *J. Geophys.*
Research, 64, 1489-1499, 1959.
- Stehn Ch. E., The geology and volcanism of the
Krakatau group, *Proc. Fourth Pac. Sci. Congr.*,
Batavia, 1929.
- Verbeek, R. D. M., *Krakatau*, Batavia, 1886.
- Williams, Howell, Calderas and their origin. *Univ.*
Calif. (Berkeley), Publs. Geol. Sci., 25(6), 239-346,
1941.
- Williams, Howell, Problems and progress in vol-
canology, *Quart. J. Geol. Sci.*, 109, 311-332, 1954.

(Manuscript received April 24, 1961;
revised June 28, 1961.)

Theoretical Phase Velocities for a Lunar Seismic Experiment

BRUCE A. BOLT

*Department of Geodesy and Geophysics
Cambridge, England*

Abstract. Theoretical frequency spectra S_{20} to S_{160} of spheroidal eigenvibrations for three models of the moon are presented. The derived phase velocities for Rayleigh waves, with periods between 120 and 20 seconds, are compared with phase velocities calculated using a layer approximation. The comparison demonstrates that, for the moon, the latter approximation is inadequate for waves having periods exceeding 25 to 30 seconds. The results (a) provide a suggestion that the operation of a single recorder of free lunar vibrations may provide discriminatory information on the interior of the moon and (b) provide data for the construction of such a recorder.

Introduction. In a recent publication, Press, and Neugebauer [1960] discussed the results of seismic experiments likely to be feasible in the initial lunar exploration. In particular, they discussed the types of data to be expected from the continuous operation of a single seismic detector. They also presented two possible lunar structures, Rayleigh dispersion curves calculated without regard to the sphericity of the moon, and attention to the possibility of inferring the properties of its interior from long waves in suitable seismographs. The results of period Rayleigh waves are worked out in detail in the present paper, which is a continuation of an earlier investigation [Bolt, 1960], and the possible range of periods of oscillations of the moon.

Interest in the results arises from experiments in both the U.S.S.R. and U.S.A. to conduct geophysical experiments on the lunar surface by means of instruments carried by rockets. The computed frequency spectra of spheroidal eigenvibrations provide information on the dynamical range required by seismic detectors designed for operation on the moon. In addition, they show that, for waves or free oscillations are excited on the moon, such detectors could provide data which would allow considerable discrimination of plausible lunar structures.

The method of computation of eigenvibrations

and phase velocities described by Bolt and Dorman [1961] was used to obtain the results presented here. Lunar models, all taken as spherical, were defined by setting down distributions of density ρ , compressional velocity α , and shear velocity β as functions of the radius r . The gravitational attraction at any interior point is then $4\pi\gamma r\rho/3$. The frequency spectra of spheroidal eigenvibrations were located by solving a determinantal equation whose elements are solutions of a sixth-order differential equation. These vibrations may be treated as the superposition of two sets of traveling waves; the phase velocity of either wave train (neglecting lunar rigid rotations) is then $c = 2\pi Rf/(l + 1/2)$, where R is the radius of the sphere, f is the eigenfrequency, and l the order of oscillation.

The assumed lunar models. Astronomical observations on the moon's figure give approximately a mean radius of 1738 km and a mean density of 3.33 g/cm³; values for the differences between the principal moments of inertia are found to be unequal [Jeffreys, 1952, p. 154]. The low mean density precludes the presence of much material as dense as iron, even after allowance for an outer shell similar to the earth's crust. The marked ellipticity, which implies a nonhydrostatic state amounting to a stress difference of about 20 atmospheres at the center, presents a difficulty for theories requiring extensive liquid regions within the moon.

Jeffreys [1937a] has considered a lunar model

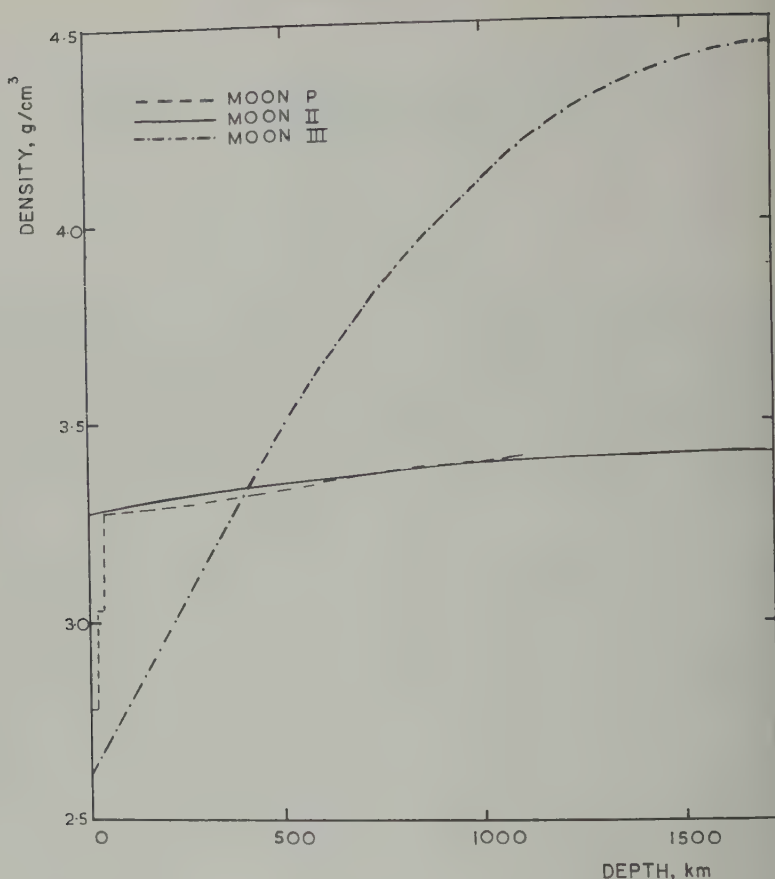


Fig. 1. Density distributions in the three lunar models considered. Calculations which involved moon model P did not require detailed knowledge of the density distribution below 1000-km depth.

entirely composed of a material similar to ultrabasic rock, with surface density 3.28 g/cm^3 . In this case self-compression alone increases the density to about 3.41 g/cm^3 at the center.

Consequences of some other speculations on the structure of the moon's interior have been considered lately in some detail. The models studied belong to two main classes. In the first class, the density does not depend strictly on the radius alone. For example, Urey [1960b] has suggested a nonuniform density distribution which varies with latitude, being least near the axis of the figure directed toward the earth. Models of the second class have density distributions with spherical symmetry. Homogeneous solid models are discussed by Jeffreys [1952] and Ramsey [1948] and models with liquid cores or near-surface regions of low rigidity by

Kuiper [1959], MacDonald [1961], and Urey [1960a].

A further possible variation is suggested by the evidence for a widespread layer of relatively low seismic velocities ('the low velocity layer' deeper than 50 km in the earth [e.g. Gutenberg, 1959; Lehmann, 1960; Bolt and Dorman, 1961]). The maximum central pressure likely for the moon is of the order of 10^5 atmospheres. (The lunar models, moon II and moon III, defined below have maximum pressures of 3×10^4 and 6×10^4 atmospheres, respectively.) This pressure value is reached at a depth about 300 km in the earth, so that pressures associated with the suggested terrestrial low velocity layer occur toward the center of the moon. No direct evidence on the interior lunar temperature is available, but temperatures at

TABLE 1. Values of θ

	Moon II	Moon III
	1.2	8.6
	1.1	15.9
	1.1	17.7

ose likely in the upper mantle of the possible [MacDonald, 1959].

nar models considered in this paper, on P, moon II, and moon III, belong second class and are assumed to be solid (to transmit shear waves) throughout, first-order discontinuities in physical properties below 20-km depth. They each satisfy mean mass and mean radius of the models and their interior properties are based on the correlation between density and seismic velocities observed for the earth. Adjustments were made in the models to obtain more realistic cases where possible after direct information on the typical surface rocks on the moon and the flow at the lunar surface becomes available. The density distributions defining the models are shown in Figure 1.

For moon II, $\rho = 3.415 - 0.135x^2$ and for moon III, $\rho = 4.430 - 1.830x^2$, where $x = r/R$. Moon P has a similar density distribution

to moon II, with a modification for a crust similar to that of the earth. Apart from a superficial layer, 1 km thick, moon P is the same as model II constructed by Press, Buwalda, and Neugebauer [1960]. The results for moon P for periods greater than about 15 sec may therefore be compared directly with the corresponding calculations of these authors.

The variations of compressional and shear velocities with depth in models II and III are derived from the density distributions by the use of the correlation laws,

$$\alpha = -2.40 + 3.12\rho \quad (1)$$

and

$$\beta = -0.60 + 1.52\rho \quad (2)$$

valid only in the range $2.5 < \rho < 4.5$, where α and β are in km/sec and ρ is in g/cm³. Equations 1 and 2 are based mainly on the velocity calculations of Jeffreys [1937b] for the earth and the density values of earth model A [Bullen, 1953]. As reported in a recent important paper, Birch [1961] obtained $\alpha = -2.55 + 3.31 \rho$ from measurements of compressional velocities in rocks at pressures up to 10^4 atmospheres. He showed, moreover, that this equation is probably consistent with a constitution of the earth's mantle resembling average

TABLE 2. Rayleigh Wave Theoretical Periods and Velocities, Moon Model P*

Plane Layers			Spherical Shells				
Phase Velocity, km/sec	Group Velocity, km/sec	No. of Layers	Mode Number	Period, sec	Phase Velocity, km/sec	Group Velocity, km/sec	No. of Layers
4.16	4.05	25	20	118.8	4.48		39
4.15	4.04	25	30	82.3	4.35		26
4.13	4.02	25	40	63.1	4.27		20
4.12	4.00	25	50	51.4	4.21		16
4.10	3.96	25	60	43.5	4.15	3.97	13
4.07	3.90	25	80	33.5	4.05	3.90	10
4.02	3.77	25	90	30.2	3.99		9
3.92	3.51	25	100	27.6	3.94		8
3.82	3.31	25	150	19.8	3.67		5

surface parameters:

h , km	α , km/sec	β , km/sec	ρ , g/cm ³
20	6.03	3.53	2.78
20	7.02	4.08	3.03

TABLE 3. Rayleigh Wave Theoretical Periods and Velocities, Moon Model II*

Plane Layers				Spherical Shells				N
Period, sec	Phase Velocity, km/sec	Group Velocity, km/sec	No. of Layers	Mode Number	Period, sec	Phase Velocity, km/sec	Group Velocity, km/sec	
100.0	4.09	4.05	25	20	121.5	4.39		
90.0	4.08	4.05	25	30	83.6	4.29		
80.0	4.08	4.05	25	40	63.7	4.23		
70.0	4.07	4.05	25	50	51.5	4.20		
60.0	4.07	4.05	25	60	43.2	4.18	4.11	
50.0	4.07	4.05	25	80	32.7	4.15	4.11	
40.0	4.06	4.05	25	90	29.2	4.14		
30.0	4.06	4.04	25	100	26.3	4.13		
20.0		150	17.7	4.10		

* Near-surface parameters:

h , km	α , km/sec	β , km/sec	ρ , g/cm ³
20	7.85	4.38	3.28
20	7.87	4.39	3.29

chondrites or mafic igneous rock. The differences between (1) and Birch's equation do not exceed about 5 per cent and are not significant in this problem.

The velocities used for moon P are those given by Press, Buwalda, and Neugebauer for their model II. Their values never differ by

more than 5 per cent from the corresponding velocities derived from (1) and (2).

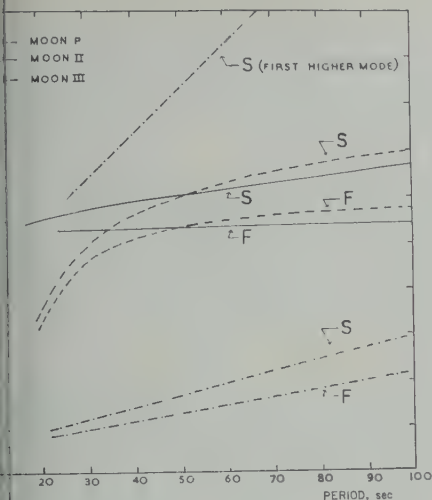
Moon II and moon III are extreme cases of lunar models which have continuous variation in α , β , and ρ throughout. Moon II has a surface density of $\rho_0 = 3.28$, near that of dunite or eclogite; for moon III, $\rho_0 = 2.60$, which

TABLE 4. Rayleigh Wave Theoretical Periods and Velocities, Moon Model III*

Plane Layers				Spherical Shells				N
Period, sec	Phase Velocity, km/sec	Group Velocity, km/sec	No. of Layers	Mode Number	Period, sec	Phase Velocity, km/sec	Group Velocity, km/sec	
100.0	3.44	3.11	25	20	139.6	3.82		
90.0	3.40	3.11	25	30	99.7	3.59		
80.0	3.37	3.10	25	40	77.6	3.48	3.13	
70.0	3.33	3.10	25	50	63.5	3.41	3.13	
60.0	3.30	3.10	25	60	53.8	3.36		
50.0	3.26	3.10	25	80	41.2	3.29		
40.0	3.23	3.09	25	90	36.9	3.27		
30.0	3.19	3.09	25	100	33.4	3.25		
20.0	3.14	...	25	150	22.7	3.20		

* Near-surface parameters:

h , km	α , km/sec	β , km/sec	ρ , g/cm ³
20	5.80	3.40	2.62
20	5.90	3.44	2.66



Rayleigh wave dispersion curves for the three lunar models. The symbol S denotes a curve calculated using the exact formulation for a gravitating sphere; F denotes a curve calculated from a plane layer approximation.

the density of granite. It is of interest to note the departures from chemical homogeneity of the two models. For this purpose we use the index θ [Bolt, 1957] so that

$$\theta = (\alpha^2 - 4/3\beta^2) d\rho/d\rho \quad (3)$$

is the pressure at radius r . Where the material is homogeneous $\theta = 1$, and departures from homogeneity normally increase θ . The values of θ at the surface, at $x = 1/2$, and at the center for the two models, are listed in Table 1.

It follows from Table 1 that the distributions of parameters chosen to define moon II yield a model with a nearly homogeneous composition. In contrast, moon III represents, as expected, a model of extreme inhomogeneity throughout the interior with accumulation of heavy materials at depth.

Phase velocity solutions. The period range for Rayleigh waves considered in this paper is 20 to 100 sec; these periods correspond to wavelengths between about 50 and 550 km. It is shown below that velocities for shorter periods may be obtained satisfactorily by using a plane layer approximation assuming plane layers. The reason detailed superficial layers, which would affect the short-period surface wave velocities, have not been included in the models

considered here. The indicated period range corresponds to the eigenfrequency spectrum for spheroidal oscillations $\omega_{S_{150}}$ to $\omega_{S_{20}}$. ($\omega_{S_i^m}$ denotes a vibration with n radial nodal surfaces and a surface pattern of displacement defined by the surface spherical harmonic P_i^m .) The periods of the various oscillations and the phase velocities deduced from them are listed for the three models in Tables 2, 3, and 4. The computations of spheroidal oscillations were made, using homogeneous spherical shells, each 20 km thick, in the way described in detail in Bolt and Dorman [1961]. The parameters α , β , ρ adopted for the two upper shells in each model are tabulated in the tables together with the number of shells N used for the estimation of each eigenfrequency.

The extreme lunar models, moon II and moon III, have markedly different periods of oscillation in almost all modes. From a previous paper [Bolt, 1960], the values of ω_{S_2} for these models are 15.1 and 12.1 minutes, respectively. The conclusion, therefore, is that the greater concentration of mass toward the center in moon III entails a relatively large reduction in period for the lower modes and a corresponding increase in period for the higher modes.

Phase velocities for Rayleigh waves in the three spherical gravitating lunar models were derived from the spheroidal eigenfrequencies and are given in Tables 2, 3, and 4. The corresponding dispersion curves are plotted in Figure 2. We note the following points.

(a) For $T > 50$ sec, the dispersion curves for moon P and moon II lie closely together, while for $T < 50$ sec the phase velocities for moon P decrease much more rapidly. This decrease is a consequence of the relatively low velocity crustal layers possessed by moon P.

(b) The dispersion curve for moon III is strikingly lower than that for moon II. It is clear that Rayleigh surface wave trains recorded by a long-period vertical seismograph on the moon would provide strong evidence for a distinction between such competing models.

The phase velocities computed by Press, Burvalda, and Neugebauer [1960] extend over the period range considered here. They remark that their method of calculation, which assumes homogeneous plane layers, is an adequate approximation only at periods short enough to be

unaffected by the sphericity of the moon. The error of the plane layer approximation is brought out quantitatively by separate computations of dispersion curves for moon P, moon II, and moon III, using a matrix method developed for automatic computing by J. Dorman (personal communications). In each case 25 homogeneous plane layers, each 20 km thick, were used; the seismic parameters adopted in each layer were exactly those taken in the corresponding spherical shell. The results are also listed in Tables 2, 3, and 4, and a graphical comparison between the three pairs of phase velocity solutions is made in Figure 2.

As has already been shown for the earth [e.g. Bolt and Dorman, 1961], the total effect of sphericity and gravity is to increase substantially at long periods the values of phase velocity above those for a correspondingly layered half-space. For the earth, at $T = 100$ sec, the phase velocity difference between spherical and flat models is about 0.05 km/sec, whereas for the moon, with its greater curvature, at $T = 100$ sec the difference is as great as 0.2 km/sec.

The conclusion from the comparison is that, in the case of the moon, the plane layer approximation deviates from the exact solution by more than about 1 per cent (~ 0.04 km/sec in phase velocity) for periods greater than 25 to 30 sec.

There is much recent evidence from observational seismology that, at least for paths across less heterogeneous parts of the earth's crust, wave trains corresponding to higher-mode solutions of the surface wave equation are clearly recorded. It may be that the lack of oceanic-continental contrast on the moon will permit widespread recording of higher-mode waves on lunar seismographs. The calculated overtones ${}_1S_i$ for moon III were used to obtain the phase velocity curve for the first higher Rayleigh mode shown in Figure 2. In this model the period of ${}_1S_{100}$ is 26 sec compared with 33 sec for ${}_1S_{100}$.

Some values for the group velocities of Rayleigh waves are given in the tables, both from the exact solution and from the plane-layer approximation. The main conclusion from a comparison of the two sets of velocities is simi-

lar to that for the corresponding earth case [Bolt and Dorman, 1961]: for $20 < T < 100$ sec corresponding group velocities differ by less than 1 per cent.

Acknowledgments. I wish to thank Dr. James Dorman for making available his machine program for the calculation of surface wave dispersion data for plane-layered media. Computing facilities were kindly made available by Lamont Observatory, Columbia University, New York, and by the United Kingdom Atomic Energy Authority.

REFERENCES

- Birch, F., Composition of the earth's mantle, *Geophys. J. Roy. Astron. Soc.*, **4**, 295, 1961.
- Bolt, B. A., Spheroidal oscillations of the moon, *Nature*, **188**, 1176, 1960.
- Bolt, B. A., Earth models with continuous density distribution, *Monthly Notices Roy. Astron. Soc. Geophys. Suppl.*, **7**, 360, 1957.
- Bolt, B. A., and J. Dorman, Phase and group velocities of Rayleigh waves in a spherical gravitating earth, *J. Geophys. Research*, **66**, 1961.
- Bullen, K. E., *An Introduction to the Theory of Seismology*, Cambridge University Press, 21, 1953.
- Gutenberg, B., *Physics of the Earth's Interior*, Academic Press, New York, chap. 4, 1959.
- Jeffreys, H., The density distributions in the inner planets, *Monthly Notices Roy. Astron. Soc. Geophys. Suppl.*, **4**, 62, 1937a.
- Jeffreys, H., The times of P, S and SKS and the velocities of P and S, *Monthly Notices Roy. Astron. Soc., Geophys. Suppl.*, **4**, 548, 1937b.
- Jeffreys, H., *The Earth*, 3rd ed., Cambridge University Press, 1952.
- Kuiper, G. P., *Vistas in Astronautics*, vol. 2, Pergamon Press, Oxford, 273, 1959.
- Lehmann, I., Structure of the upper mantle as derived from the travel times of seismic P and S waves, *Nature*, **186**, 956, 1960.
- MacDonald, G. J. F., Calculations on the thermal history of the earth, *J. Geophys. Research*, **66**, 1967, 1959.
- MacDonald, G. J. F., Interior of the moon, *Science*, **133**, 1045, 1961.
- Press, F., P. Buwalda, and M. Neugebauer, A lunar seismic experiment, *J. Geophys. Research*, **66**, 3097, 1960.
- Ramsey, W. H., On the constitution of the terrestrial planets, *Monthly Notices Roy. Astron. Soc.*, **108**, 406, 1948.
- Urey, H. C., Criticism of the melted moon theory, *J. Geophys. Research*, **65**, 358, 1960a.
- Urey, H. C., The origin and nature of the moon, *Endeavour*, **19**, 87, 1960b.

(Manuscript received July 5, 1961.)

Record of Cosmic-Ray Intensity in the Meteorites

J. R. ARNOLD AND M. HONDA

*School of Science and Engineering, University of California, San Diego
La Jolla, California*

D. LAL

*Tata Institute for Fundamental Research
Bombay, India*

Abstract. Extensive data now exist on the abundance of radioactive and stable nuclides produced by cosmic rays in iron meteorites. Half-lives of radioactive species range from 16 to 1.2×10^9 years. In this paper these data are compared with calculated production rates based on derived energy spectra of nuclear-active particles in meteorites and experimental excitation functions. Both relative and absolute amounts of the various species are in approximate agreement if the flux of cosmic rays is assumed not to have varied. The variation, averaged over the half-life of each species, does not appear to exceed a factor of 2 in any case, except possibly K^{40} , for which the data are still insufficient. It is concluded that the cosmic-ray intensity has been constant, in this sense, at least over millions of years.

INTRODUCTION

Production of measurable quantities of He isotopes by cosmic-ray bombardment was first predicted by Bauer [1948] and Ley [1948]. The observation by Paneth, and Mayne [1952] of a high ratio of He^4 in meteoritic He confirmed this prediction. Then a voluminous experimental and theoretical literature has developed in this

meteorites are targets containing a record of cosmic-ray bombardment to which they have been subjected. Thus it should be possible to obtain from them information on the past intensity of the cosmic radiation [Geiss, 1957]. Availability of new data on radioactive products [Signer and Nier, 1960; Honda and Ald, 1961; and data of E. I. Fireman, Signer, and H. Wänke] has led us to attempt a new analysis of this problem.

The first step is to calculate the relative and absolute production rates of the observed nuclides in meteorites using the present-day cosmic-ray flux. If this flux has been constant in the past, the rates of half-life short compared with the time of bombardment are in secular equilibrium. Observed rates of decay should equal calculated production rates. Disagreement would indicate a variation in cosmic-ray intensity.

The radioactive nuclide K^{40} does not appear to be in secular equilibrium [Voshage and Hintenberger, 1959; Honda, 1959]. A comparison of the concentrations of K^{40} with those of stable and shorter-lived species gives us information on the variation of the cosmic-ray flux over the time of bombardment.

This paper will deal with iron meteorites. Stone meteorites will not be considered here since their complex chemical composition makes calculation of production rates more difficult.

METHOD

A number of methods have been used for calculation of production rates. All employ experimental information on the cosmic radiation. Cascade theory [Martin, 1953; Ebert and Wänke, 1957; Hoffman and Nier, 1958; Signer and Nier, 1960], the distribution of star sizes in emulsion [Shedlovsky, 1960; Goel, in press], and the laboratory thick-target bombardment data [Fireman and Zähringer, 1957] are also employed. All these different techniques should yield equivalent results, given the necessary data.

We have chosen to use derived spectra of primary and secondary cosmic-ray particles at 100 g/cm² and 10 g/cm² depth in a meteorite. These are combined with excitation functions to yield production rates. This method appears

to us to be the most accurate for prediction of production rates at present. The simulated cosmic-ray bombardment technique of *Fireman and Zähringer* [1957], however, may offer more promise for the future, if the laboratory bombardment conditions are sufficiently realistic.

In meteorites below 500 kg in weight, the change in production rate with depth is moderately small. Thus, in Grant (440 kg), the decrease in the concentration of neon isotopes from edge to center is 40 per cent [*Signer and Nier*, 1960]. Other rare gases show a smaller change. The effect of size in the region from 20 to 500 kg (8.5 to 25 cm radius) on production rate is also small. This may be seen from a comparison of data on Mn^{55} , Al^{26} , and Be^{10} in Grant, Williamstown, and Aroos [*Honda and Arnold*, 1961; *Honda, Shedlovsky, and Arnold*, 1961], and also from data on Ar^{39} in chondrites [*Stoerner, Schaeffer and Davis*, 1960].

The choice of a standard depth of 100 g/cm² (13 cm) in a meteorite is a convenient one. Many data are available for this depth in the atmosphere. As is shown below, these can be related to the meteorite case. From a point at this depth in a meteorite of radius 200 g/cm², the distribution of distance to the surface with angle, plotted (for example) against cosine θ from 1 to -1, is closely similar to the distribution in a very large object on the scale of cosine θ from 1 to 0 (upper hemisphere). Thus the cascade development should be much the same. We may take our reference point, then, at 100 g/cm² depth in a meteorite of radius 200 g/cm² (mass 550 kg). The results will not be very sensitive to this choice. We will also calculate the spectrum at 10 g/cm² depth.

Because of the uncertainty in the amount of atmospheric ablation, the calculations made at a given depth in the preatmospheric meteorite apply to a smaller, unknown depth in the present object. This difficulty is not very serious in the present estimates of ablation (2 to 15 cm) but is realistic. Conversely, experimental data on the concentration of spallation products cannot be used to make accurate estimates of ablation.

The differential energy spectrum of nuclear active particles at a moderate depth in an atmosphere or meteorite is made up of contributions of primary and secondary particles, protons, neutrons, and mesons (α particles are present in comparatively small numbers). We may conclude from both experiment and theory that the energy spectrum is continuous and monotonically decreasing in the region of interest. The qualitative importance of this fact may be seen in Table 1. Here the cross section for the production of each species at 730 Mev is compared with the observed decay rate in the meteorite Aroos. At both ends of the table the ratio of decay rate to cross section is high, dropping to a minimum between. For species in the Mn region, the cross section becomes high below 100 Mev. Here the importance of the low-energy particles is unmistakable. The ratio drops rather smoothly until the point is reached at which most of the production takes place in the region of 730 Mev. For products of lower Z the cross section at 730 Mev drops off, and the ratio begins to rise as production moves steadily to higher energies.

We now proceed to derive the relative energy spectra. They will then be multiplied by the best available excitation functions to give the relative production rates for radioactive and stable

TABLE 1. Content of Radioactive Nuclides in Aroos Compared with Production Cross Section in Bombardment of Fe with 730-Mev Protons

Left Column			Right Column		
Content in Aroos, dpm/kg	σ (exp.) 730 Mev, mb	Ratio	Content in Aroos, dpm/kg	σ (exp.) 730 Mev, mb	Ratio
Be ¹⁰	4.1	1.0*	Ca ⁴⁵	5.	1.4*
Na ²²	2.1	0.36†	Sc ⁴⁶	30.	6.4*
Al ²⁶	3.6	0.43†	V ⁴⁹	164.	30.*
Sj ³²	0.8	0.3*	Cr ⁵¹	260.	27.*
Cl ³⁶	14.	6.8†	Mn ⁵⁴	470.	33.†
Ti ⁴⁴	4.4	2.*			

* M. Honda and D. Lal, to be published.
† *Honda and Lal* [1960].

The spectra at 10 and 100 g/cm² depth normalized using the flux of high-energy and the total energy flux. This provides absolute production rates. Finally, any of the cosmic-ray intensity with considered.

ENERGY SPECTRUM

The composition and the energy of primary cosmic rays are rather uniform. We made use of the data obtained by McDonald [1959] using Cerenkov-scintillation detectors. Observations during periods of abnormally high solar activity in 1959 have shown that the flux of particles above 100 Mev kinetic energy per nucleon was considerably lowered. However, the spectra for 1956 may be taken as typical. The solar activity has a large effect on the low-energy particles but a much smaller effect on the total energy flux, which mainly determines the flux of secondary particles at moderate depth. The preferential spectrum of protons, α particles, and heavier nuclei can be fairly well represented by $(1 + E)^{-2.5}$ for E , the kinetic energy per nucleon, greater than 1.5 bev. Above a given energy the proportion of protons, α particles, O group, and heavier nuclei has been found to be $10^8 : 1.5 \times 10^8 : 13.3 : 3.9$ [McDonald, 1959; Rao, Biswas, Daniel, Neelaiah, and Peters, 1958]. At lower energies the spectrum falls off. The exact form is given by McDonald for protons and α particles above 350 Mev per nucleon. We neglect particles below this energy.

The average energy of a cosmic-ray nucleon is about 100 Mev. Using this spectrum, to be 4 bev. The flux of secondary particles in planetary space near the earth can be obtained from McDonald's data and satellite observations [Vernov and Chudakov, 1959; Van der Kamp and Frank, 1959]. The latter include some data on secondary particles. The total flux is 2460 particles/cm² sec sterad (2100 protons) or 4000 nucleons/cm² sec sterad. The corresponding values for the omnidirectional (4π) flux are 3.1 and 6.0 particles/cm² sec. The omnidirectional energy flux is about 1000 Mev/cm² sec.

For simplicity we shall divide the discussion of the secondary energy spectra into three parts: greater than 3 bev, 100 Mev to 3 bev, and 2 Mev to 100 Mev.

The most straightforward region is that above 3 bev. A large number of experiments have shown that in the average nuclear interaction, up to tens of bev, a single nucleon carries away most of the incident energy [Bogachev, Buniatov, Merekov, Sidorov, and Yarba, 1960; Daniel, Rao, Malhotra, and Tsuzuki, 1960; Kalbach, Lord, and Tsao, 1960]. Pi mesons carry 25 to 40 per cent of the energy, depending on the size of the target nucleus. Low-energy nucleons share 0.3 to 0.4 bev almost independent of the primary energy. The loss of a moderate, constant, fraction of energy in each collision has the effect of leaving the spectral shape nearly unaltered. The near constancy of the relative spectrum in this energy region should continue to depths much greater than 100 g/cm². Even the charged π mesons in this region are produced with the same spectrum as that observed for primary particles [Powell, Fowler, and Perkins, 1959]. Thus the total particle spectrum above 3 bev retains closely the form of the primary energy spectrum.

Next, we consider the secondary particles in the range of energy 0.1 to 3 bev. Particles in this energy region are mostly nucleons emerging from the cascade in the target nucleus. A fraction are charged mesons. The energy spectrum of charged particles emerging from disintegrations produced by cosmic rays has been studied by the Bristol group [Powell, Fowler, and Perkins, 1959] in nuclear emulsions exposed at high altitudes. The number-energy spectra have been given for all important types of charged particles. They are found not to depend very strongly on latitude and altitude of exposure. We assume that the neutron energy spectrum above 100 Mev is identical with the measured production spectrum of protons. It also seems safe to assume that the production spectrum of nucleons in emulsion nuclei (average $Z = 30$) holds good for interactions produced in iron.

Charged particles will be rendered ineffective for nuclear reaction if they are brought to rest by ionization. The range-energy relation in iron is known [Atkinson, 1957], and we have calculated the probability of nuclear interaction as a function of energy for each type of charged particle. Thus we have estimated the production spectrum of effective particles. Since the

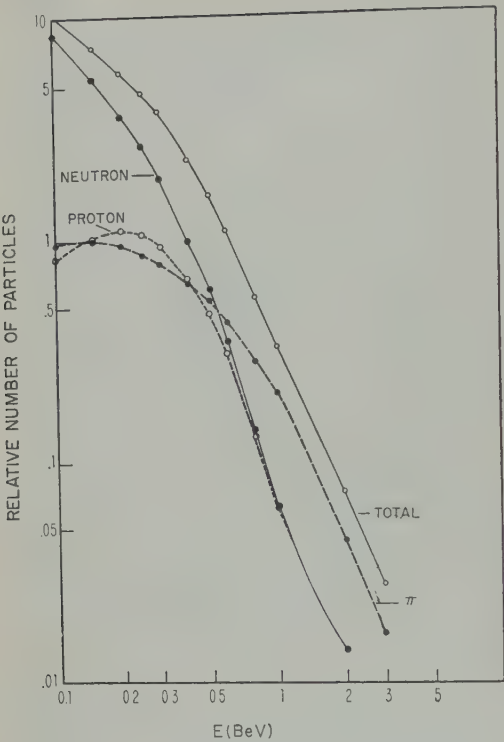


Fig. 1. Differential flux of nuclear-active particles at a depth of 100 g/cm² in a meteorite, as deduced from data given by Powell. The contributions of protons, neutrons, and charged π mesons are shown separately along with the total.

data of Powell and co-workers refer to nuclear interactions well inside the atmosphere, they include those produced by low-energy secondary particles. The slow variation with altitude and latitude which they find is expected to hold good so long as the spectrum of particles above 1 bev (which are mainly responsible for the generation of secondaries) remains unchanged. Further down in the atmosphere low-energy particles increase in number. The spectrum be-

comes steeper, especially in the region of 0.1-1 bev. At greater depths the spectrum approaches a steady state, as is seen in several experiments [Birnbaum, Shapiro, Stiller, O'Dell, 1952; Soberman, 1956; Hess, Patterson, Wallace, and Chupp, 1959].

We may obtain the spectrum of particles in iron from that in emulsion in the atmosphere using the important experiment of Shapiro, Stiller, Birnbaum, and O'Dell [1951]. A block of lead roughly equivalent to a sphere of 15-cm radius was carried to a level of 22 g/cm² in the atmosphere at geomagnetic latitude 56°. Small stacks of emulsion were placed at various points within the block. The star size distribution in these emulsions was compared with that in free air. The total star production rate was two times higher in the block, but the size distribution was unchanged within experimental error. Passing through the block, the total rate and the distribution underwent small changes only. The changes would be still smaller in outer space, where the incoming flux covers 4 π rather than 2 π radians, although a small 'skin' effect in the outer few centimeters due to low-energy primaries would appear.

This result justifies our use of the spectra derived for emulsion in the atmosphere to deduce the energy spectra of particles in an iron block.

The sum of the flux spectra of protons, neutrons, and π^+ mesons constitutes the total spectrum of nuclear-active particles. The data of Marquez [1952] and others show that reactions of protons and neutrons have similar cross sections in this region. The propriety of lumping in π^+ mesons with nucleons is discussed in Appendix 1.

The individual spectra, and their sum, are shown in Figure 1. The total spectrum is well represented by the expression $(0.2 + E)^{-2.5}$ from 0.1 to 3 bev.

TABLE 2. Energy Spectra of Nuclear-Active Particles *

Symbol	Remarks	0.002-0.1 bev	0.1-3 bev	>3 bev
$S(100, E)$	100 g/cm ² depth	$60(1 + 0.01E^{-1} + 1.1 \times 10^{-5}E^{-2})$	$3.3(0.2 + E)^{-2.5}$	$5.8(1 + E)^{-2.5}$
$S(10, E)$	10 g/cm ² depth	$37(1 + 0.01E^{-1} + 1.1 \times 10^{-5}E^{-2})$	$7.2(0.4 + E)^{-2.5}$	$11(1 + E)^{-2.5}$
$S(pr, E)$	Primary spectrum		†	†

* Units are particles/cm² sec bev.
† McDonald [1959].

TABLE 3. Cross Sections Assumed for Normalization of Excitation Functions

mb	Refer- ence	mb	Refer- ence
3 Bev			
0	<i>f, k, n, o, p, q</i>	Be ⁷	13 <i>a, c, d</i>
0	<i>e, f, s</i>	Be ¹⁰	7 <i>b</i>
0	<i>e, f, s</i>	C ¹⁴	3
1 Bev			
1.6	<i>e, f, r</i>	K ⁴¹	16 <i>r</i>
1.6	<i>e, f, r</i>	Ca ⁴¹	7
1.6	<i>e, f, r</i>	Ca ⁴²	18 <i>r</i>
0.70	<i>a, d, g, h, i</i>	Ca ⁴³	20 <i>r</i>
0.76	<i>a</i>	Ca ⁴⁴	22 <i>r</i>
0.4	<i>b</i>	Ca ⁴⁵	1.5 <i>b, h</i>
3.1	<i>d, g</i>	Ca ⁴⁶	0.2
2.2	<i>d, g</i>	Sc ⁴⁵	25 <i>r</i>
7.5	<i>a</i>	Sc ⁴⁶	7.0 <i>b, h</i>
9.0	<i>e, f, r</i>	Ti ⁴⁴	2 <i>b</i>
5.0	<i>e, f, k</i>	V ⁴⁸	25 <i>h</i>
2	<i>e, f, r</i>	V ⁴⁹	30 <i>b, h</i>
6.4	<i>e, f</i>	V ⁵⁰	14
2.5		Cr ⁵¹	25 <i>b, h</i>
0.06	<i>b, j</i>	Mn ⁵²	10 <i>a, h, m</i>
3.	<i>r</i>	Mn ⁵³	30
10.	<i>b</i>	Mn ⁵⁴	30 <i>a</i>
		Fe ⁵⁵	50 <i>l</i>

and Lal [1960].
 and D. Lal, to be published.
 Friedlander and Hudis [1958].
 [1957].
 and Zähringer [1959].
 and Zähringer, in press.
 and Miller, Wolfgang, Hudis, and Baker
 m, Stevenson, and Folger [1952].
 Miller, and Seaborg [1951].
 er, Schaeffer, and Davis [1960].
 an and Zähringer [1957].
 witz, Rowland, and Friedlander [1958].
 r and Markowitz, unpublished.
 s, Libby, and Wolfgang [1956].
 [1959].
 [1958].
 an [1955].
 Appendix 3, formula A5.
 ringer, private communication.

work on the spectral shape can be made by
 (a) the observed star size distribution
 emulsions exposed in the atmosphere
 altitudes [Birnbaum, Shapiro, Stiller,
 ell, 1952], and in the lead block of
 Stiller, Birnbaum, and O'Dell [1951],
 the star size distribution in emulsions
 by artificially accelerated particles of

different energies. This procedure is described
 in Appendix 2. The agreement is good.

Lastly, we discuss the energy spectrum of
 particles from 2 to 100 Mev. This region is
 important for nuclides close to the target nucleus
 and for small fragments. Only neutrons occur in
 important numbers. Information can be obtained
 from the experimental-theoretical neutron spec-
 trum of Hess, Patterson, Wallace, and Chupp
 [1959], and from the Bristol data [Powell, Fowler,
 and Perkins, 1959]. The Bristol measurements
 refer to protons, and cannot be used below 20
 Mev because of Coulomb barrier effects. The
 results of Hess apply to light nuclei in the
 atmosphere. Their applicability to meteorites
 cannot be proved by reference to the experi-
 ment of Shapiro, Stiller, Birnbaum, and O'Dell
 [1951], since the stars studied by them are
 produced almost entirely by particles above
 100 Mev. The measurements of Gross [1956] on
 the energy spectrum (1 to 10 Mev) of neutrons
 arising from disintegrations in C, Al, Ni, Ag, and
 Au indicate that above 3 Mev no appreciable
 differences exist in the neutron production
 spectra. Below 3 Mev the spectrum is steeper
 for carbon than for other targets. The Bristol
 and Hess curves agree above 20 Mev. As the
 best approximation available at present we
 adopt the Hess spectrum from 2 to 100 Mev.
 The shape of this spectrum is well represented by
 the expression $k(1 + 0.01E^{-1} + 1.1 \times 10^{-6}E^{-2})$.
 The final spectrum is given in Table 2 as

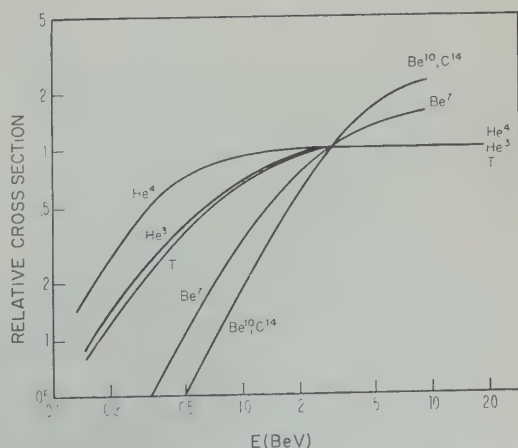


Fig. 2. Assumed cross sections for the production
 of small nuclei from iron, normalized at 3 bev.

TABLE 4. Assumed Cross Sections* below 100 Mev, in mb

Nuclide	E(Mev)												Target Nuclide
	2	2.8	4	5.7	8	11	16	22	31	44	62	87	
H ³							2.5	5	5	5	5	5	Fe ⁵⁶
He ³							5	10	10	10	10	10	Fe ⁵⁶
He ⁴							50	100	100	100	100	100	Fe ⁵⁶
V ⁴⁸											1	8	Fe ⁵⁶
V ⁴⁹											4	10	Fe ⁵⁶
V ⁵⁰											4	6	Fe ⁵⁶
Cr ⁵¹									17	67	30	30	Fe ⁵⁶
Mn ⁵²									6	9	9	9	Fe ⁵⁶ , Fe ⁵⁴
Mn ⁵³							9	36	34	48	90	55	Fe ⁵⁶ , Fe ⁵⁴
Mn ⁵⁴							3	2	40	140	65	45	Fe ⁵⁶ , Fe ⁵⁴
Fe ⁵⁵			1.5	3	5	6	10	770	580	340	230	150	Fe ⁵⁶
Fe ⁶⁰									0.1	0.3	0.2	0.1	Ni ⁶⁴ , Ni ⁶²
Co ⁵⁶							1	3	13	20	9	6	Ni ⁶⁰ , Co ⁵⁹
Co ⁵⁷							8	10	10	7	6	4	Ni ⁶⁰ , Co ⁵⁹
Co ⁵⁸	4	7	6	10	13	16	19	23	18	11	5	3	Ni ⁶⁰ , Co ⁵⁹ , Ni ⁶⁴
Co ⁶⁰					1	2	2	1					Ni ⁶⁰
Ni ⁵⁹							1	3	3	2	1	1	Ni ⁶⁰

* These are cumulative cross sections. Neutron capture processes are not included. The cross sections for production are based on total Fe. For the Co and Ni isotopes they are based on total Fe, assuming 0.43 per cent Ni and 0.43 per cent Co [Honda and Arnold, 1961]. The Fe⁵⁴ (n, p) Mn⁵⁴ excitation function is taken as that of Fe⁵⁶ (n, p) Mn⁵⁶ [Hughes and Schwartz, 1958], lowered by 3 Mev.

S(100, E). The three regions are joined smoothly; the absolute normalization is described below.

The spectrum at 10 g depth, S(10, E), has also been derived. It is given in Table 2. Between 0.1 and 3 bev it has been calculated from the analysis of star size distributions mentioned above. At lower energies the curve of Hess is assumed to be valid.

EXCITATION FUNCTIONS

Excitation functions for the various product species are now required. We shall divide these into two regions: above and below 100 Mev.

Above 100 Mev, for A < 15, we have normalized our excitation functions at 3 bev. The values chosen are given in Table 3. Up to Be⁷ experimental data are available, both for the normalization and for the shape of the excitation function. For Be¹⁰ the shape has been estimated, and the measured cross section at 0.73 bev has been used to fix the scale. The shape for C¹⁴ has been assumed the same as for Be¹⁰; the cross section at 3 bev has been estimated according to Barr [1957]. The normalized excitation functions are shown in Figure 2.

From A = 20 to A = 54 the shape of the

excitation function is calculated from

$\ln \sigma(E, A, Z) = \ln P - P \Delta A + C(A, Z)$ (see Appendix 3). This equation is fairly consistent with the shape of the excitation function above 0.1 bev where data are available. We have normalized each curve at 1 bev, using experimental data for proton bombardment of Fe where available, data on targets of similar Z (often Cu), formula A5 in Appendix 3 for cumulative cross sections, or where necessary the Rudstam equation [Rudstam, 1956; Honda and Lal, 1960]. The values used for normalization are shown in Table 3. For Fe⁵⁶ the Cu⁶³ (p, p) Cu⁶² curve has been used [Markowitz, Rowland, and Friedlander, 1958].

Below 100 Mev the bombarding particles are mainly neutrons. Here very few experimental data are available. The cross section for H¹ is certainly large down nearly to the (n, α) threshold, and for H² and He³ it may be. The values assumed are arbitrary. For A \geq 48 we have used the excellent data of Sharp, Diamond, and Wilkinson [1956] for protons on Co⁵⁹. The cross section for Fe⁵⁶ (n, xp, yn) is assumed equal

TABLE 5. Production Rates Relative to Cl^{36}

Aroos Expt*	$S(100)$	$S(10)$	$S(pr)$	Nuclide	Aroos Expt*	$S(100)$	$S(10)$	$S(pr)$
	12	10	8	Ca^{41}		1.1	1.0	0.8
34†	27	23	17	Ca^{42}		3.1	2.6	1.9
128†	135	108	63	Ca^{43}	3.9¶	3.6	3.0	2.0
	0.69	0.74	1.0	Ca^{44}		4.2	3.5	2.1
0.29	0.31	0.34	0.62	Ca^{46}	0.3	0.28	0.22	0.13
	0.13	0.15	0.27	Ca^{48}	0.09¶	0.05	0.04	0.02
1.2†	0.73	0.81	1.3	Sc^{46}		5.0	4.0	2.3
0.15	0.11	0.12	0.18	Sc^{48}	2.1	1.5	1.2	0.63
0.26	0.10	0.11	0.15	Ti^{44}	0.31	0.38	0.31	0.19
0.06	0.05	0.06	0.06	V^{48}	6	7.3	5.2	2.2
	0.41	0.44	0.49	V^{49}	12	9.6	6.7	2.6
	0.30	0.32	0.35	V^{50}	10¶	5.2	3.5	1.2
(1)	(1)	(1)	(1)	Cr^{51}	20	19	11	2.0
(1.2)†	1.2	1.2	1.2	Mn^{52}		6.6	4.2	0.8
1.3†	0.65	0.65	0.61	Mn^{53}	38	33	18	2.5
1.9†	1.6	1.5	1.4	Mn^{54}	34	38	20	2.5
0.9-1.0†	0.90	0.85	0.74	Fe^{55}	120§	220	100	6.2
	0.38	0.35	0.31	Fe^{60}		0.1	0.1	0.03
	0.11	0.09	0.08	Co^{56}	(4.4)	4.5	2.1	0.1
	1.8	1.7	1.5	Co^{57}	6.4	5.5	3.0	0.6
0.6	1.5	1.4	1.2	Co^{58}	(4.4)	17	7.8	0.1
	2.6	2.3	1.8	Co^{60}	1.3	0.6	0.3	0.01
				Ni^{59}	3.4	1.4	0.7	0.1

and Arnold [1961], unless otherwise noted.

gas data are from Signer and Nier [1961]

an and DeFelice [1960]; Wänke, unpublished; Heymann and Schaeffer [1961].

ty 1600 ± 600 dpm/kg from M. Honda, unpublished.

lated from the value 7.8 ± 0.5 dpm/kg obtained by Stauffer and Honda, [1961].

er and Honda [1961] and unpublished data.

of Co^{59} (p, xp, yn). This procedure is
 rary but systematic. Sharp and co-
 nherent evidence that it is reliable to
 factor of 2, except where the product
 gic number of neutrons or protons,
 experimental cross section is low. This
 o Mn^{53} , but no correction has been
 e data of Ashby, Catron, Newkirk, and
 [1958] on Fe^{56} ($n, 2n$) Fe^{55} , and data of
 and Schwartz [1958] on Cu^{63} ($n, 2n$) Cu^{62}
 n used to obtain the excitation function

imation of excitation functions for Fe^{60}
 and Ni isotopes is discussed in Appendix

RELATIVE PRODUCTION RATES

relative production of each species is

$$Z) = \int \sigma(E, A, Z) S(X, E) dE$$

$$= \int \sigma(E, A, Z) [ES(X, E)] d \ln E \quad (2)$$

For convenience, $\ln E$ has been used as the independent variable in the actual numerical integrations. The results are given in Table 5. These are normalized to $\text{Cl}^{36} = 1$. The yields are cumulative where appropriate (e.g., Ar^{36} includes Cl^{36} ; Ar^{40} includes Cl^{40} but not K^{40}). The experimental data for the radioactive species in the iron meteorite Aroos [Honda and Arnold, 1961] are given for comparison. The species Be^{10} , Al^{26} , and Mn^{53} have the same abundance in Grant and Williamstown as in Aroos within 20 per cent or less [Honda, Shedlovsky, and Arnold, 1961]. Values for the rare-gas isotopes in Aroos, normalized to $\text{Ar}^{36} = 1.2$, are also given.

The contribution of particles below 100 Mev to the production of each species is given in Table 6. In extreme cases such as Co^{58} and Fe^{56} this contribution is of the order of 90 per cent of the total.

TABLE 6. Production by Particles below 100 Mev
(Relative to total Cl^{36})

Nuclide	$S(100^*)$	Nuclide	$S(100)$
H^3	1	Mn^{53}	20
He^3	2	Mn^{54}	25
He^4	18	Fe^{56}	190
V^{48}	0.9	Fe^{60}	0.01
V^{49}	1.3	Co^{56}	3.9
V^{50}	0.8	Co^{57}	3.1
Cr^{51}	10	Co^{58}	17
Mn^{52}	2.5	Co^{60}	0.6
		Ni^{59}	1.0

* To obtain relative production by particles below 100 Mev for $S(100)$ multiply by 0.44.

In addition to the calculations given in Table 5, we have made some estimates of the production rates at depths of 500 to 1000 g/cm² in a large meteorite. We estimate that the ratio $\Sigma \text{Ne}/\text{Ar}^{36}$ will drop about 30 per cent below the value for 100 g/cm², and that $\text{Sc}^{46}/\text{Ar}^{36}$ will be about 30 per cent higher. Such species as Mn^{53} , however, produced mainly by low-energy neutrons, will increase relative to Ar^{36} by factors of 3 to 5.

A further comparison can be made with results of Wänke [1960] on stable cosmogenic Sc^{45} . He finds the ratio $\text{Sc}^{45}/\text{Ne}^{22}$ nearly constant at 18–20 in several meteorites. From this the experimental ratio $\text{Sc}^{45}/\Sigma \text{Ne}$ may be given as 6–7.

ABSOLUTE ENERGY SPECTRA AND PRODUCTION RATES

The energy spectra at 100 and 10 g/cm² depth can be normalized in a number of different ways. Two methods appear to us most reliable for this purpose: using the flux of particles of $E \geq 3$ bev, and using the total energy flux.

First we must consider the effect of complex nuclei. The mean free paths of these nuclei are smaller than that of protons, and the general effect of nuclear collision is that the bombarding nucleus breaks up (complete breakup of heavier nuclei will take more than one collision). Below 50 g/cm² depth in the meteorite we assume that the flux is the same as it would be if all the primary nucleons entered the meteorite as free particles. The effect of this assumption is to overestimate the flux somewhat, but no data are available for a correction. At a depth of

10 g/cm² the error introduced by this assumption is somewhat larger, but since it is of minor significance only for particles coming from the hemisphere the difference will not be great.

The fact that one nucleon carries away a major fraction of the energy in collisions in the bev region is discussed under 'Energy Spectrum' above. The inelasticity, defined as the energy carried by π mesons in the laboratory system, is around 0.3 for air and 0.5 for iron [Babakov, Grigorov, Dubrovín, Mischenkov, Muratov, Sarycheva, Sobiniakiv, and Rapoport, 1959]. The probability of emitting a π meson of energy greater than 3 bev is appreciable only above about 20 bev incident nucleon energy. Above 0.4 bev is carried by evaporation and low-energy cascade nucleons. Since the primary flux falls rapidly with energy the situation is simplified for distances of one or two mean free paths. We need only consider the production of fast nucleons by the primaries. The rate of decrease of the flux of nucleons of $E > 3$ bev can be calculated on the reasonable assumption that the emerging nucleon moves in the forward hemisphere relative to the direction of the primary. λ being defined as the mean free path for nucleon interaction, and λ_{abs} as the mean free path for decrease of the high-energy flux, it can be shown that

$$1 - (\lambda_{\text{int}}/\lambda_{\text{abs}}) = (1 - K)^S \quad (1)$$

where K = inelasticity and S = slope of the integral primary spectrum. For iron, λ_{abs} is approximately equal to $1.5 \lambda_{\text{int}}$. The uncertainty in this value is less than 20 per cent. The interaction mean free path in iron is 100 g/cm² which corresponds to $\lambda_{\text{abs}} = 150$ g/cm². Integration over angle for our standard meteorite at 100 g/cm² depth yields the result that the flux of particles of $E \geq 3$ bev is reduced to 0.25 of the original value. An error in the value of λ_{abs} can be considered as an error in the coordinates of our assumed standard depth.

The coefficient of $S(100)$ above 3 bev is determined from this result and the primary spectrum. The coefficients in the lower energy ranges are determined by the condition of continuity.

Integration of the spectrum $S(100, E)$ using this normalization yields the result that 60 per cent of the total energy present in the primary beam is retained at this depth. The remaining

ent corresponds to losses in the form of n , π^0 mesons, and escape of particles from meteorite.

approaches have been taken to obtain independent estimate of the energy loss to depth. First we consider the differences in total energy carried by the particles in secondary spectrum given by the Bristol data and that of the effective spectrum obtained after correction for ionization losses, π^0 escape, and minor effects. An approximate value of 45 per cent is obtained for the loss. Another estimate may be based on discussion of energy balance in the atmosphere given by *Puppi* [1956]. A correction is made for the fact that π^+ mesons, which do not interact in the atmosphere, are usually present in the meteorite. By estimating energy loss by this procedure, a value of 40 per cent is obtained. Both estimates agree well with the results obtained above.

total number of nuclear-active particles per unit depth may also be derived. The value obtained is 16 particles per incident primary. Of these about 80 per cent are neutrons below 10 g/cm². This figure seems reasonable on the basis of the Bristol data [*Powell, Fowler, and Fowler* 1959].

spectrum at 10 g/cm² may also be normalized in the same way. In a meteorite of thickness 5 cm the calculated reduction in the number of high-energy particles (integrated over 4π) is about 50 per cent.

rate of production, in atoms/min kg, is

$$\frac{\times 10^{23} \times 10^3 \times 60}{56 \times 10^{27}} \int S \sigma dE$$

$$= 0.65 \int S \sigma dE \quad (4)$$

is expressed in millibarns. The absolute production rate of Cl^{36} for $S(100)$ is 23 dpm/kg. For $S(10)$ and $S(pr)$ the rates are calculated to be 35 dpm/kg, respectively. The absolute production for all other species in Table 1 is obtained by multiplying by these factors. The value of 23 dpm/kg Cl^{36} may be compared with the experimental figure of 14 dpm/kg in the difference, a factor of 0.6, may well be due to the combined error of the absolute normalization and the excitation function.

TABLE 7. Observed and Calculated Depth Effect in Grant*
(Limiting case of zero ablation)
 $C(10)/C(100)$

Nuclide	Observed†	Calculated
He ³	1.19	1.3
He ⁴	1.08	1.2
ΣNe	1.32	1.6
Ar ³⁶	1.28	1.5
Ar ³⁸	1.23	1.4

* *Signer and Nier* [1960].

† Sample S-71a, bar N, is compared to sample S-47, bar J.

It is also possible that the result would be better if the effect of atmospheric ablation of the meteorite could be included. The important conclusion is that an absolute flux equal to or lower than the experimental one is sufficient to produce all the species considered.

We may now consider the depth effect in the production of rare gases. The change in production of rare gases with depth in Grant (440 kg present mass) may be compared with calculated values, for the limiting case of zero atmospheric ablation. This is done in Table 7, using data obtained by *Signer and Nier* [1960] for two points at approximately 10 and 100 g/cm² present depth. It appears that some atmospheric ablation has taken place, although the amount is very difficult to estimate.

DISCUSSION

Study of Table 5 shows that the agreement between the calculated and observed values for the relative production rates of various stable and radioactive species is very good. The average deviation for each species (except Al^{26} and Co and Ni isotopes) is about 30 per cent. The high observed values of Al^{26} are presumably due to the bombardment of P and S in the meteorites. The trends in production rates are accurately reproduced. In particular the factor of about 1000 between the production rates of Fe^{65} and Na^{22} is seen to come about because of the abundance of low-energy particles in the secondary spectrum. Although depth effects are not explored in detail, the agreement is good, allowing for some ablation. It should be noted that our model predicts a sharp increase with depth for $A \geq 53$, increases persisting down to $A = 48$.

By far the largest discrepancy is that for Co^{58} . Even if the entire observed yield of Co^{56+58} is assigned to this species, the predicted value is high by a factor of 3. The production of this species is mainly by $\text{Ni}^{58}(n, p)\text{Co}^{58}$. For this type of reaction nearly all production occurs in the region of the peak, below 50 Mev. At high energies the cross section becomes very small. The experimental value of 560 ± 110 mb at 14 Mev [Purser and Titterton, 1958] has been used to normalize the $\text{Fe}^{56}(n, p)\text{Mn}^{56}$ cross-section curve, which peaks at 14 Mev. There is also a low-energy contribution, and one from $\text{Co}^{59}(n, 2n)\text{Co}^{58}$. The calculation for Fe^{56} deviates by a smaller amount in the same direction. Perhaps the flux at energies ~ 14 Mev has been overestimated. The species Co^{60} and Ni^{59} can also be produced by the (n, γ) process [Van Dilla, Arnold, and Anderson, 1960]. This production has not been included in Table 5. It will be discussed in a separate paper.

The success of the model in predicting decay rates for the species of short half-life is evidence for its applicability.

Its equal success in predicting the decay rates of the long-lived species Be^{10} , Al^{26} , Si^{32} , Cl^{36} , Ti^{44} , and Mn^{53} is strong evidence of the constancy of the cosmic-ray flux over the last few million years. More precisely, we consider the average quantity.

$$\int_0^T \int_0^\infty S(E, X, t) \sigma(E, A, Z) e^{-\lambda(A, Z)t} dE dt$$

where t is reckoned positive backward from the time of fall, and T is the time when bombardment commenced. This integral appears to have the same value, within at most a factor of 2, for each radioactive product species considered. Such a statement does not preclude the possibility of short, intense bursts of cosmic radiation (or solar radiation), provided that the integral is not much affected.

Any variation in the cosmic-ray bombardment over periods of the order of 10^8 to 10^9 years can only be studied using cosmogenic K^{40} . This may be done by comparing bombardment ages obtained using K^{40} and a stable species with those obtained with a shorter-lived radioactive nuclide and a stable one. The pairs $\text{K}^{40}\text{--K}^{41}$ and $\text{Cl}^{36}\text{--Ar}^{36}$ are good examples. The same age should be obtained from both pairs, if the bombardment intensity has been constant.

If the K^{40} content is relatively lower, and if the $\text{K}^{40}\text{--K}^{41}$ age longer, the level of bombardment was lower in the past, and the true duration longer than that derived from such pairs $\text{C}^{32}\text{--Ar}^{36}$. As in the case of Williamstown [Honda, Shedlovsky, and Arnold, 1961], the question of terrestrial age of the meteorite must be considered if the meteorite is a 'find.'

At present the data are fragmentary. A direct measurement of the ratio $\text{K}^{40}/\text{K}^{41}$ by Voshage and Hintenberger [1959] yielded a result that Carbo is 0.4 to 1.3×10^9 years older than Treysa. Other pairs have given 5×10^8 and 7 to 12×10^8 years for the ages of Treysa and Carbo [Sprenkel, 1959; Vilček and Wänke, 1961] respectively.

According to Honda [1959] the absolute concentrations of cosmogenic K^{40} in Grant and Williamstown are consistent with the bombardment ages of about 5×10^8 years obtained from other pairs [Schaeffer and Zähringer, 1961; Heymann and Schaeffer, 1961; Honda, Shedlovsky, and Arnold, 1961]. For Aroos, ages of 5.9 to 5.4×10^8 years have been reported [Heymann and Schaeffer, 1961; Vilček and Wänke, 1961] using $\text{Cl}^{36}\text{--Ar}^{36}$ and $\text{Cl}^{36}\text{--Ne}$. The value of 2.3 dpm/kg reported by Honda and Arnold [1961] is definitely lower than predicted for these ages. Recently Stauffer and Honda, [1961] have made a new mass-spectrometric determination, using K^{39} carrier. They obtained 4.9×10^{-10} g/g or 7.8 ± 0.5 dpm/kg. The older result must be attributed to loss of K, at an extremely low concentration, in the chemical processing. The present value is consistent with the predictions of Table 5.

These facts suggest, but do not yet prove, that even over the past 0.5 to 1×10^9 years the level of cosmic-ray intensity has been similar to the present one.

The observed constancy of the cosmic-ray radiation intensity over various periods up to millions of years is to be expected on the basis of current theories of cosmic-ray origin [Ginzburg, 1959; Hayakawa, Ito, and Terashima, 1958; Hayakawa and Koshida, 1959]. Theoretical predictions concerning variations on a 10^9 -year time scale are not easy to make. In the theory of Shklovsky and Ginzburg, type I supernovae play an essential role. Since these occur in our galaxy at the rate of about one every hundred years, and since the life of each cosmic-ray particle in the galaxy

longer, the fluctuations should be averaged. However, a change in the rate of production of cosmic rays by type I supernovae in the past 10^9 years is quite possible. Similar considerations apply to other theories, since there is an appreciable fraction of the life of the galaxy.

APPENDIX 1. π MESONS

There are differences in the excitation functions of isotopes for π mesons and protons. The literature suggests that in the bevel region the size distribution of π -produced stars is very closely that of proton stars. At higher energies, however, the situation is not quite clear. Analyses of π -produced stars in nuclear emulsions have been made at 35–50, 120, 220, and 280 and 750 Mev [Blau and Gerasimovic, 1956; Ivanova, Ostroumov, and Pavlova, 1956; Sprague, Haskin, Glasser, and Schein, 1954; Bernardini and Levy, 1951]. The results seem to indicate a somewhat steeper size distribution than for protons of kinetic energy equal to the total π^- energy. At low energies, the only results on π^+ stars are in the 35–80 Mev [Sprague, Haskin, Glasser, and Schein, 1954]. Here we find that the π^+ size distribution is flatter than that of protons. If we take an equal proportion of π^+ and π^- in the meteorite, the resulting star size distribution seems to be very similar to that of proton stars. Qualitatively, we conclude that the excitation functions for combined π mesons and protons are similar.

We have not considered the reactions induced by π capture at rest. Radiochemical analyses have been made in iodine [Winsberg, 1954] and in silver [Sugihara and Libby, 1952]. This method, when used with nuclear emulsion data, would indicate that about 30 per cent of stars are pronged (emission of neutrons only). The size distribution of stars is similar to that of stars induced by 200 Mev protons. The capture events can therefore be considered equivalent to a flux of 200 Mev protons. The average number of slow π^- mesons produced in iron is less than 15 per cent [Bernardini and Levy, 1951]. This is equivalent to raising the flux of neutrons in the energy region 100 to 200 Mev by less than 20 per cent. We neglect the effect.

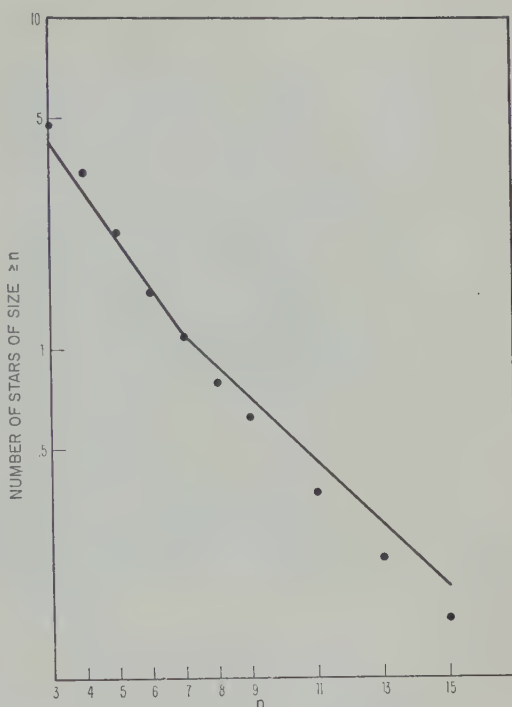


Fig. 3. Comparison of experimental and calculated relative integral star size distributions in emulsion at 100 g/cm² depth. The line gives smoothed experimental data; points are calculated.

APPENDIX 2. STAR SIZE DISTRIBUTIONS

Star size distributions for incident beams of monoenergetic protons have been determined at numerous points between 0.1 and 6 beV, for number of prongs $n \geq 3$ [Lock, March, Muirhead, and Rosser, 1955; Morand, Baudinet-Robinet, and Winand, 1958]. Morand and co-workers have fitted the experimental distributions with an empirical equation. With the aid of this equation we have obtained the integral star size distributions at energies throughout the region of interest, spaced at a constant logarithmic interval. A crude but adequate estimate of $\sigma_s(E)$, the total cross section for production of stars of three or more prongs, is made using the data of Germain [1951] for stars of $n \geq 2$, and correcting for three-pronged stars. σ_s drops to a low value below 100 Mev and approaches constancy at beV energies.

For a given differential energy spectrum of star-producing radiation $S(X, E)$, the integral star size distribution may be calculated from

$$T(n) = \int I(E, n) \sigma(E) S(X, E) dE$$

$$= \int I(E, n) \sigma(E) [ES(X, E)] d \ln E \quad (A1)$$

This distribution may be compared with the experimental one, or the function S can be varied until a good fit is obtained.

The function $T(n)$ as given by Birnbaum, Shapiro, Stiller, and O'Dell [1952] for a depth of 50 g/cm² in the atmosphere is plotted in Figure 3 along with the function derived from the distribution $S(100, E)$. The distribution $S(10, E)$ in the region 0.1 to 3 bev was derived using equation A1.

APPENDIX 3. SHAPE OF EXCITATION FUNCTION ABOVE 100 Mev

We assume that the total inelastic scattering cross section is constant throughout the region of interest, and that

$$\sum_{28}^{56} \sigma(A) = \sigma_{in}$$

We further assume that

$$\ln \sigma(E, A) = PA - q(E) \quad (A2)$$

and that the distribution of $\sigma(Z)$ among isobars is independent of E . From these assumptions we may conclude, for any E ,

$$\sum_{A=28}^{56} e^{PA-q} = e^{-q} \sum_{A=28}^{56} e^{PA}$$

$$= \frac{e^{56P} - e^{27P}}{e^q(1 - e^{-P})} = \sigma_{in} \quad (A3)$$

We approximate by setting $(1 - e^{-P}) = P$ and neglecting e^{27P} , a procedure that does not introduce large errors in the region of interest. Then

$$\ln \sigma_{in} = 56P - q - \ln P \quad (A4)$$

Eliminating q ,

$$\ln \sigma(E, A) = \ln P - P \Delta A + \ln \sigma_{in} \quad (A5)$$

where $\Delta A = 56 - A$, while

$$\ln \sigma(E, A, Z)$$

$$= \ln P - P \Delta A + C(A, Z) \quad (A6)$$

We have derived $P = 0.11E^{-0.64}$ from published data for protons of 0.34 to 6 bev on Cu, using $R = 1.8$ and $S = 0.469$ for spallation products in the region $|Z - SA| \leq 1.5$ [Rudstam, 1956]. A similar relation was used earlier for Fe [Honda and Lal, 1960].

The approximations used (and the assumption of constant σ_{in} throughout the energy region) introduce appreciable errors. However, simple equation A6 fits the available data well as any more exact expression we have used. We believe it to be adequate from $A = 28$ to $A = 54$. It has the great advantage of reducing the subjective element in the selection of experimental cross-section data.

APPENDIX 4. Ni AND Co CROSS SECTIONS

The cross sections for Ni and Co isotopes are sums of cross sections for production from stable Ni and Co species. Processes like (p, xn) , (π^+, x) on Fe are unimportant. All cross sections from Co⁵⁹ were assumed equal to those for production of the same species by protons [Sharp, Diamond, and Wilkinson, 1956]. These were normalized to an abundance of 0.4 per cent. Co. Ni⁵⁸ (n, p) Co⁵⁸ was taken from experiments below 4 Mev and at 14 Mev [Purser and Titterton, 1958]. Ni⁶⁰ (n, p) Co⁶⁰ was taken as the same. Fe⁵⁶ (n, p) Mn⁵⁶ [Hughes and Schwartz, 1956]. Co⁶⁷ from Ni⁵⁸ was taken as twice Ni⁵⁸ (n, p) Ni⁵⁷ [Hughes and Schwartz, 1958] at low energy or twice Cu⁶³ (p, pn) [Markowitz, Rowland, and Friedlander, 1958] at high. Co⁵⁶ from Ni⁵⁸ was assumed equal to twice Mn⁵⁶ from Fe⁵⁶, because of cumulative yield. Cross sections for Co⁵⁸, Co⁵⁹, and Ni⁵⁹ were estimated similarly. For Co⁶⁰ and Fe⁶⁰ higher isotopes of Ni were also considered.

Acknowledgments. This research was supported by the Office of Ordnance Research, U. S. Army and the National Aeronautics and Space Administration. We are grateful to R. R. Daniel, R. A. Subramaniam, and many other colleagues for helpful discussions, and to Margaret Casady for assistance in the calculations.

REFERENCES

- Ashby, V. J., H. C. Catron, L. L. Newkirk, and C. J. Taylor, *Phys. Rev.*, **111**, 616, 1958.
- Atkinson, J. H., High energy particle data, UCRL-2426, 1957.
- Babayan, C. P., N. L. Grigorov, M. M. Dubrov, L. G. Mischenkov, V. S. Murzin, L. I. Sarychev, V. S. Sobiniakiv, and I. D. Rapoport, *Proc. Moscow Cosmic Ray Conf.*, **1**, 178, 1959.
- Baker, E., G. Friedlander, and J. Hudis, *Phys. Rev.*, **112**, 1319, 1958.
- Barr, D. L., *AEC Rept. UCRL-3793*, 1957.
- Batzel, R. E., D. R. Miller, and G. T. Seaborg, *Phys. Rev.*, **84**, 671, 1951.
- Bauer, C. A., *Phys. Rev.*, **72**, 354, 1947; **74**, 225, 1948.
- Bernardini, G., and F. Levy, *Phys. Rev.*, **84**, 61, 1951.

- Bull. Am. Phys. Soc., [2]3, 221, 1958.
- n, M., M. M. Shapiro, B. Stiller, and F. Dell, *Phys. Rev.*, **86**, 86, 1952.
- , and A. K. Oliver, *Phys. Rev.*, **102**, 489,
- y, N. P., S. A. Buniatov, V. P. Merekov, Sidorov, and V. A. Yarba, *J. Exptl. Theophys. USSR*, **38**(4), 1346, 1960.
- y, A., *Phys. Rev.*, **114**, 878, 1959.
- y, A., W. F. Libby, and R. L. Wolfgang, *Rev.*, **101**, 1557, 1956.
- y, R., N. Kameswara Rao, P. K. Malhotra, Y. Tsuzuki, *Nuovo cimento*, **16**, 1, 1960.
- H., and H. Wänke, *Z. Naturforsch.*, **12a**, 57.
- E. L., *Phys. Rev.*, **97**, 1303, 1955.
- E. L., and J. DeFelice, *J. Geophys. Res.*, **65**, 3035, 1960.
- E. L., and J. Zähringer, *Phys. Rev.*, **107**, 957.
- er, G., J. M. Miller, R. Wolfgang, J. and E. Baker, *Phys. Rev.*, **94**, 727, 1954.
- Chimia (Switz.)*, **11**, 349, 1957.
- E., *Phys. Rev.*, **82**, 596, 1951.
- V. I., *Progress in Elementary Particle Cosmic Ray Physics*, vol. 4, chapter 5, 1958.
- K., and J. Zähringer, *in press*.
- S., *Geochim. et Cosmochim. Acta*, *in press*.
- , The absolute yield of low-energy neutron from 190 Mev proton bombardment of silver, nickel, aluminum, and carbon, *The-RL-3330*, 1956.
- a, S., K. Ito, and Y. Terashima, *Progr. Theophys. Suppl.* **6**, 1, 1958.
- a, S., and M. Koshiba, *Progr. Theoret. Phys.*, **21**, 473, 1959.
- N., H. W. Patterson, R. Wallace, and Chupp, *Phys. Rev.*, **116**, 445, 1959.
- , D., and O. A. Schaeffer, *Abstracts*, 42nd Meeting Am. Geophys. Union, Washington, April 1961.
- J. H., and A. O. Nier, *Phys. Rev.*, **112**, 958.
- I., *Geochim. et Cosmochim. Acta*, **17**, 148,
- M., and J. R. Arnold, *Geochim. et Cosmochim. Acta*, **23**, 219, 1961.
- M., and D. Lal, *Phys. Rev.*, **118**, 1618,
- M., J. P. Shedlovsky, and J. R. Arnold, *in. et Cosmochim. Acta*, **22**, 133, 1961.
- D. J., and R. B. Schwartz, Neutron cross *sections*, AEC Rept. BNL 325, 2d edition, 1958.
- H. E., *Nature*, **161**, 356, 1948.
- N. S., V. I. Ostrolov, and Yu. V. *Soviet Phys. JETP*, **37**, 1137, 1961.
- R. M., J. J. Lord, and C. H. Tsao, *Phys. Rev.*, **113**, 325, 1960.
- O., P. V. March, H. Muirhead, and W. Rosser, *Proc. Roy. Soc. London*, **A**, **230**, 55.
- z, S., F. S. Rowland, and G. Friedlander, *Phys. Rev.*, **112**, 1295, 1958.
- Marquez, L., *Phys. Rev.*, **88**, 225, 1952.
- Martin, G. R., *Geochim. et Cosmochim. Acta*, **3**, 288, 1953.
- McDonald, F. B., *Phys. Rev.*, **116**, 462, 1959.
- Morand, M., Y. Baudinet-Robinet, and L. Winand, *Proc. 2nd Intern. Conf. Geneva*, **30**, 109, 1958.
- Paneth, F. A., P. Reasbeck, and K. I. Mayne, *Geochim. et Cosmochim. Acta*, **2**, 300, 1952.
- Powell, C. F., P. H. Fowler, and D. H. Perkins, *The Study of Elementary Particles by the Photographic Method*, Pergamon Press, New York, p. 442, 1959.
- Puppi, G., The energy balance of cosmic radiation, *Progr. in Cosmic Ray Phys.*, **3**, 1956.
- Purser, K. H., and E. W. Titterton, *Rept. ANU/P-200*, Australian National University, 1958.
- Rao, M. V. K. A., S. Biswas, R. R. Daniel, K. A. Neelakantan, and B. Peters, *Phys. Rev.*, **110**, 751, 1958.
- Rudstam, G., Ph.D. thesis, Uppsala, 1956.
- Rudstam, G., P. C. Stevenson, and R. L. Folger, *Phys. Rev.*, **87**, 358, 1952.
- Schaeffer, O. A., and J. Zähringer, *Z. Naturforsch.*, **13a**, 346, 1958.
- Schaeffer, O. A., and J. Zähringer, *Phys. Rev.*, **113**, 674, 1959.
- Schaeffer, O. A., and J. Zähringer, *Geochim. et Cosmochim. Acta*, **19**, 94, 1960.
- Shapiro, M. M., B. Stiller, M. Birnbaum, and F. W. O'Dell, *Phys. Rev.*, **83**, 455, 1951.
- Sharp, R. A., R. M. Diamond, and G. Wilkinson, *Phys. Rev.*, **101**, 1493, 1956.
- Shedlovsky, J. P., Ph.D. thesis, Princeton University, 1960.
- Signer, P., and A. O. Nier, *J. Geophys. Research*, **65**, 2947, 1960.
- Soberman, R. K., *Phys. Rev.*, **102**, 1399, 1956.
- Sprague, A. D., D. M. Haskin, R. G. Glasser, and M. Schein, *Phys. Rev.*, **94**, 994, 1954.
- Sprenkel, E., Ph.D. thesis, University of Rochester, 1959.
- Stauffer, H., M. Honda, *J. Geophys. Research*, **66**, 3584-3586, 1961.
- Stoener, R. W., O. A. Schaeffer, and R. Davis, Jr., *AEC Rept. BNL-4862*, 1960.
- Sugihara, T. T., and W. F. Libby, *Phys. Rev.*, **88**, 587, 1959.
- Van Allen, J. A., and L. A. Frank, *Nature*, **184**, 219, 1959.
- Van Dilla, M. A., J. R. Arnold, and E. C. Anderson, *Geochim. et Cosmochim. Acta*, **20**, 115, 1960.
- Vernov, S. N., and A. Ye. Chudakov, *Doklady Akad. Nauk SSSR*, **125**, 304, 1959 [translated in *Soviet Phys. Doklady*, **4**, 338, 1959].
- Vilesek, E., and H. Wänke, *Z. Naturforsch.*, **16a**, 379, 1961.
- Voshage, H., and H. Hintenberger, *Z. Naturforsch.*, **14a**, 828, 1959.
- Wänke, H., *Z. Naturforsch.*, **15a**, 953, 1960.
- Winsberg, L., *Phys. Rev.*, **95**, 198, 1954.

Cosmic-Ray-Induced Radioactivity in Terrestrial Materials

RAMA¹ AND MASATAKE HONDA

*School of Science and Engineering, University of California, San Diego
La Jolla, California*

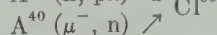
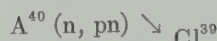
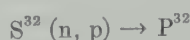
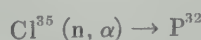
Abstract. Cosmic-ray-induced activities have been detected in a few laboratory materials both at sea level and at mountain altitude. The measured amounts are found to be consistent with those expected to be produced from known flux of cosmic-ray neutrons and negative muons.

Introduction. While passing through the atmosphere cosmic rays interact with air and produce easily detectable quantities of radioactivity. As they reach the earth's surface their intensity, and consequently the induced radioactivity, becomes very small. This radioactivity can, however, be measured by using sensitive techniques of low-level

measurements. The major part of the radioactivity in various materials at sea level is produced by neutrons and negative muons. The muons are the decay products of pions which, in turn, are secondaries of interactions caused by cosmic rays in the earth's atmosphere. Neutrons, like pions, are direct secondaries of such interactions. At sea level the total flux of neutrons is comparable to that of negative muons. Neutrons interact strongly with matter and thus cause reactions which may result in the production of radionuclides. Muons, on the other hand, are weakly interacting particles which undergo only a few disintegrations by direct collision. The disintegrations caused by capture of negative muons, however, may not be entirely negligible. When a negative muon comes to rest in a material it falls rapidly into the *K* shell of an atom. For a muon this shell is very close to the nucleus, and there is a finite probability that the muon may get captured in the nucleus and decay. This probability depends on the atomic number of the arresting material. The values for some of the materials are: H(27%), C(9%), N(17%), O(27%), F(37%), Al(61%), S(76%), Fe(92%), Pb(96%), and U(96%) [Sens, 1959;

Tennet, 1960]. The capture occurs through the elementary reaction $\mu^- + p \rightarrow n + \nu$. The neutrino takes away most of the energy, but enough excitation remains with the nucleus to enable it to emit a few neutrons. In a small percentage of the cases, charged particles are also emitted. The end products, in some cases, may be radioactive and sufficient for detection.

We have not yet made any attempt to detect the radioactivity in any terrestrial material taken directly from its natural environment. Instead, we have measured the radioactivity in some laboratory materials exposed to cosmic rays at sea level and at mountain altitude. The materials chosen were hydrochloric acid, carbon tetrachloride, sulphuric acid, sulphur dioxide, argon, and mercury. The radioactivities sought were due to isotopes P^{32} (14.5 d, 1.7 Mev β^-) in HCl, CCl_4 , H_2SO_4 , and SO_2 ; Cl^{39} (55 m, 1.9 Mev β^-), in Ar; and Au^{198} (2.7 d, 0.96 Mev β^-), and Au^{199} (3.1 d, 0.30 and 0.25 Mev β^-) in Hg. These radioactivities result mainly from the following reactions:



We have detected all these activities both at sea level and at mountain altitude and have found the observed magnitudes to be consistent with those estimated from the known flux of cosmic-ray neutrons and muons.

Experimental procedure. The experiments were conducted during the period August 1960

¹ Came from the Tata Institute of Fundamental Research, Bombay, India.

TABLE 1. Summary of Experimental Results

No.	Location	Target ^a	Weight, kg	Net cpm at the Time of First Counting	Saturation Specific Activity, dpm/kg	Product/Tail
1.	La Jolla, Calif., sea level	40% HCl	5.4	0.10 ± 0.03	0.28 ± 0.07	P ³² /Cl
2.	La Jolla, Calif., sea level	40% HCl	8.1	0.17 ± 0.02	0.22 ± 0.03	P ³² /Cl
3.	La Jolla, Calif., sea level	CCl ₄	3.6	0.18 ± 0.02	0.21 ± 0.03	P ³² /Cl
4.	La Jolla, Calif., sea level	96% H ₂ SO ₄	4.0	0.15 ± 0.02	0.46 ± 0.05	P ³² /S
5.	La Jolla, Calif., sea level	Ar	52.0	1.6 ± 0.2	0.2 ± 0.02	Cl ³⁹ /Ar
6.	La Jolla, Calif., sea level	Hg	20.4	0.53 ± 0.05	0.115 ± 0.011	Au*/Hg
7.	La Jolla, Calif., sea level	Hg	20.4	0.44 ± 0.04	0.103 ± 0.011	Au*/Hg
8.	Echo Lake, Colo., 685 g/cm ²	40% HCl	5.4	1.11 ± 0.03	2.4 ± 0.12 0.8 ± 0.4	P ³² /Cl P ³³ /Cl
9.	Echo Lake, Colo., 685 g/cm ²	96% H ₂ SO ₄	4.0	0.87 ± 0.07	5.1 ± 0.3 0.0 ± 0.8	P ³² /S P ³³ /S
10.	Echo Lake, Colo., 685 g/cm ²	SO ₂	0.55	0.17 ± 0.02	4.8 ± 0.5	P ³² /S
11.	Echo Lake, Colo., 685 g/cm ²	Hg	20.4	4.2 ± 0.2	0.81 ± 0.04	Au*/Hg
12.	La Jolla, Calif., under water (2.1 m)	Hg	34.0	0.60 ± 0.06	0.08 ± 0.01	Au*/Hg
13.	La Jolla, Calif., under ground (30 m)	Hg	34.0	0.15 ± 0.03	0.006 ± 0.003	Au*/Hg

^a Target size: HCl, H₂SO₄ and CCl₄ stored in glass bottles of 2 liters capacity; argon 13 kg/tank; 5-lb lecture bottle; Hg (thickness 4 cm) in polyethylene containers.

to January 1961. The experimental details are given in Appendix 1.

Results. The results of a number of experiments are given in Table 1. The values given for specific activities correspond to saturation exposure. Corrections have been made in cases where the exposure lacked saturation.

The activities observed for radiogold have been split into two parts: neutron induced and negative muon induced. This is done by comparing the results of the experiments conducted at sea level, at mountain altitude, and under water. Comparison is made as follows:

Let the specific activity induced by neutrons at sea level be x (dpm/kg) and that by muons be y (dpm/kg); then

$$x + y = 0.109 \pm 0.011 \quad (1)$$

For the mountain (Echo Lake) experiment,

the neutron-induced radiogold activity should be 11 times that at sea level, as can be seen from the relative activities of P³² observed at Echo Lake and at La Jolla. The corresponding factor for muon-induced activity as judged from the relative flux of slow muons [Rossi, 1948] at sea level and at an altitude of 685 g/cm², should be 4. Thus we have

$$11x + 4y = 0.81 \pm 0.04$$

From (1) and (2) we obtain $x = 0.053 \pm 0.005$ and $y = 0.056 \pm 0.015$. Similarly, for the underwater experiment we have

$$0.35x + y = 0.08 \pm 0.01$$

(Here the absorption thickness for cosmic-ray neutrons in water has been taken as 200 g/cm², the value found by Harding, Lattimore, Li, and Perkins [1949] for absorption of star-produced

TABLE 2. Cosmic-Ray-Induced Activities at Sea Level^a

Produced Nuclide	Observed Production Rate, atoms/min kg	Calculated Production Rate, atoms/min kg	Reference
Au ¹⁹⁸ + Au ¹⁹⁹	0.064 ± 0.007(neutron)	0.05 - 0.1 (n)	This work
	0.059 ± 0.011 (μ ⁻)	0.087 (μ ⁻)	
Cl ³⁹	0.2 ± 0.02	0.2 (n)	This work
		0.09 (μ ⁻)	
S ³⁵	0.076 ± 0.019		(L.A.H.)
P ³³	0.034 ± 0.008		(L.A.H.)
P ³²	0.041 ± 0.010		(L.A.H.)
P ³²	0.52 ± 0.06	1.0	This work
P ³³	0.09 ± 0.04	0.1	This work
P ³²	0.24 ± 0.03	0.38	This work
Be ⁷	0.049 ± 0.012		(L.A.H.)

neutron-induced activities observed at Echo Lake [Lal, Arnold, and Honda, 1960 (L.A.H.)] are multiplied by a factor of 11 to obtain values corresponding to sea level exposure.

of cosmic rays in ice. The change in the number of muons under 2.1 meters of water has been estimated to be negligible.)

(2) and (3) we obtain $x = 0.051 \pm 0.006$ and 0.062 ± 0.011 , in excellent agreement with the above result.

Calculation of production rates. For calculation of production rates of various activities we must know the energy spectrum of neutrons and the flux of slow muons. The neutron data of Peterson, Wallace, and Chupp [1959] and the muon data of Rossi [1948] are appropriate for the purpose. Using these, we have calculated the various production rates; the results are given in Appendix 2.

Discussion. The observed and the calculated activities for various activities at sea level are given in Table 2. The values for observed activities were taken from Table 1 and have been corrected for the effects due to absorption of neutrons in finite thickness of the target and in external materials like containers for the target and the roofing under which exposures were made. The total corrections amount to 10 per cent or less. No such corrections were made for muon-induced activities.

Lal, Arnold, and Honda [1960] have measured the production of a few spallation products of nitrogen and oxygen at Echo Lake. Their results are given in Table 2; the activities observed have been reduced by a factor of 11 to obtain values corresponding to sea level exposure. Rossi [1956] measured the activity of Cl³⁹ in water. Cl³⁹ was presumed to be produced by interactions of slow negative muons

with atmospheric argon. Because of many uncertainties involved in assessing the influence of various meteorological factors, however, it is difficult to obtain an estimate of the production rate of Cl³⁹ from his results.

From Table 2 it can be seen that the observed activities are somewhat lower than the calculated values. The discrepancies, however, are within the errors of experiment and calculations.

From Tables 1 and 2 the following facts may further be noted:

(i) At sea level, the activity of radiogold (and radiochlorine) induced by neutrons is approximately equal to that induced by negative muons.

(ii) The neutron-induced activities of P³² in various materials at sea level are lower by a factor of 11 than those observed at an altitude of 685 g/cm². This yields a value of 144 g/cm² for the absorption thickness of cosmic-ray neutrons in the atmosphere. This is in excellent agreement with the value obtained by Rossi [1952] for absorption of star-producing components of cosmic rays.

(iii) The radiogold activity observed under 30 meters of rock is not inconsistent with that expected from flux of slow muons at this depth (see Appendix 2).

(iv) The induced radioactivity in laboratory reagents, even at sea level, is not entirely negligible. Therefore, caution should be exercised in choosing reagents for an experiment in which a very small amount of radioactivity is to be extracted from a sample which requires treatment with large quantities of reagents.

The present work is useful in the sense that

we can now obtain a fairly good idea of the magnitudes of production of various radioactivities in terrestrial materials by cosmic-ray neutrons and negative muons. It may be pointed out that only a small fraction of negative muons give rise to radioactivity. In the atmosphere and in the oceans most of the negative muons undergo decay because of the small probability of nuclear capture in light materials which form the major constituents of air and sea water. These remarks apply in the case of the earth's crust also, since oxygen (a light material) is the major constituent of the crust. However, sizable amounts of radioactivity and rare gases may be produced in some concentrated sources of heavy materials in the crust. A possible case is that of Al^{26} resulting from silicon.

APPENDIX 1

Chemical separation. P^{32} : Carriers of phosphate (~ 20 mg) and iron (~ 20 mg) along with some bromine water were added to HCl , H_2SO_4 , and CCl_4 targets. HCl and H_2SO_4 were evaporated to small volume and then diluted with water. After performing a sulphide scavenge, the solutions were passed through a 10-ml column of cation exchange resin (Dowex 50), in hydrogen form. H_2S was expelled from the effluent and magnesium ammonium phosphate precipitated. The precipitate was ignited to $\text{Mg}_2\text{P}_2\text{O}_7$ and deposited for counting.

Activity was extracted from CCl_4 by solvent extraction. CCl_4 was shaken with 1 liter of dilute HCl . The aqueous layer was further treated in the manner described above.

SO_2 was bubbled through water containing an equivalent amount of ammonia and phosphate carrier. The same solution was used for cleaning the inside walls of the SO_2 tank after making it acidic with acetic acid. The solution was evaporated to dryness and ammonium salts were decomposed. The residue was dissolved in dilute HCl and treated in the manner described above for other targets.

The phosphate yields in various experiments ranged between 50 and 100 per cent.

Cl^{39} : 200 ml of distilled water and 9 mg chloride carrier (as HCl) were put into each tank. The tanks were filled with argon to a pressure of about 1600 psi. One atmosphere of air was left in. After the argon was exposed to cosmic rays for 1 day in the laboratory, the tanks

were rolled on the ground so that the walls would collect the activity from the walls. (It was presumed that all the Cl^{39} activity settled on the inside walls of the tank.) Argon was then let off and the water drained out. The water was warmed and the chloride precipitated as AgCl . (The precipitate contained also some iron which came from the tank walls.) AgCl was dissolved in ammonia and filtered. The filtrate was acidified with HNO_3 and AgCl reprecipitated. The precipitate was dried and deposited for counting. The whole operation was completed in 44 minutes.

The amount of chloride recovered was 66 mg, i.e. 30 mg in excess to the amount added. We believe that the over-all chemical yield in the experiment should be about 90 per cent.

Au^{198} and Au^{199} : Gold carrier (~ 30 mg) was added to mercury. After the exposure was completed, the mercury was distilled away. The residue from distillation contained the gold and a few grams of mercury. This was treated with 1 : 1 HNO_3 . The mercury was dissolved immediately, leaving gold residue in powder form. This was cleaned with distilled water and alcohol and deposited for counting. The recovery of gold was close to 100 per cent, except in experiment 6 in which 33 mg of gold in excess of the added amount was recovered. For this experiment, undistilled mercury was used. The mercury probably contained a small quantity (1.5 ppm) of gold. In later experiments, only the distilled mercury was used over and over again and the recovery of gold was always close to 100 per cent.

It is desirable to complete the distillation of mercury in as short a time as possible so that the corrections for the decay or build-up of radiogold activity during the time of distillation will be kept small. The distillation time in our experiments ranged between 15 and 20 hours and the corrections amounted to about 10 per cent, except in experiment 13 for which the correction was very large. In this experiment mercury was irradiated under 30 meters of rock but was distilled in the laboratory at sea level. The build-up of activity during the time of distillation, though small, amounted to two thirds of the total observed activity.

Measurement of activities. The sources were deposited on split copper or Plexiglas cylinders and their β activities measured on a cylindrical

unter. The radiochemical purity of the samples was ascertained by measuring their activity. To determine dpm from counting rates it is necessary to know the counting efficiency (cpm/dpm) in each case. The counter was regularly calibrated with standard β sources of various energies [Honda, Shedlovsky, and Chappell, 1961]. The efficiencies for thin (1.5 mg/cm²) sources of Cl^{36} , P^{32} , Au^{198} , and Au^{199} , determined from the calibration curve, were 41, 36, 22 per cent, respectively, on copper holders and were 28 and 15 per cent for P^{32} and Au^{198} , respectively, on Plexiglas holders. The counting efficiencies for actual sources were determined by further correcting for self-absorption. For estimating these corrections, values of 36, 5.5, and 5 mg/cm² were used for the self-thickness of β 's of Cl^{36} , P^{32} , Au^{198} , and P^{33} , respectively. The thickness of the sources was small and the corrections amounted to less than 25 per cent, except in the case of P^{33} for which the correction was about 50 per cent.

APPENDIX 2

Calculation of production rates. Neutron activities: The production rate of a nuclide by neutrons in a reaction can be calculated if the excitation function for the reaction is known. The data on excitation functions of many reactions are in general very incomplete, and the rates of production can therefore be determined only approximately. Rates of production of P^{32} in the reactions $\text{Cl}^{35}(\text{n}, \alpha) \text{P}^{32}$ and $\text{S}^{32}(\text{n}, \alpha) \text{P}^{32}$ can, however, be estimated fairly well where exist sufficient data on the excitation functions of these reactions.

Production of P^{32} from chlorine (75% Cl^{35} , 25% Cl^{37}) results mainly from the reaction $\text{Cl}^{35}(\text{n}, \alpha) \text{P}^{32}$. This reaction has a threshold at 1.0 Mev; the cross section rises sharply to 90 mb at 1.5 Mev and reaches 140 mb at 14 Mev (see neutron cross sections, BNL-325). The excitation function at higher energies is not known. However, in analogy to the reactions $\text{Mg}^{25}(\text{n}, \text{p}) \text{Na}^{24}$, $\text{Cl}(\text{p}, \text{spallation}) \text{P}^{32}$ and $\text{Al}^{27}(\text{n}, \text{p}) \text{Na}^{24}$ [Meadows and Holt, 1951a, b; Meadows, 1956; Paul and Clarke, 1952; Knox, 1949; Knox, 1952], one may assume that the cross section drops to a low value at ~ 20 Mev, rises to ~ 50 mb at ~ 40 Mev, and stays

constant at this value for higher energies. Using this excitation function and the neutron data of Hess, Patterson, Wallace, and Chupp [1959] we calculate the P^{32} production to be 0.36 atoms/min/kg Cl. To this, of course, should be added the contribution from the reaction $\text{Cl}^{37}(\text{n}, \alpha) \text{P}^{32}$. The excitation function for this reaction is not known, but the production can be estimated by assuming that the yield of P^{32} in this reaction is similar to that of S^{35} in the reaction $\text{Ar}^{40}(\text{n}, \alpha) \text{S}^{35}$. The latter has been measured by Lal, Arnold, and Honda [1960] by exposing argon directly to cosmic rays at Echo Lake (altitude 685 g/cm²). They find the production rate of S^{35} to be 0.84 atoms/min/kg Ar. The value for an exposure at sea level should be lower by a factor of 11, i.e. 0.08 atoms/min/kg Ar. The yield of P^{32} also should presumably be close to 0.08 atoms/min/kg Cl^{37} or 0.02 atoms/min/kg Cl. The total production rate of P^{32} , therefore, should be about $0.36 + 0.02$, i.e. 0.38 atoms/min/kg Cl.

Production rate of P^{33} from chlorine in the reactions $\text{Cl}^{35}(\text{n}, 2\text{pn}) \text{P}^{33}$ and $\text{Cl}^{37}(\text{n}, \alpha) \text{P}^{33}$ may also be roughly estimated. The yields of these reactions are expected to be considerably lower than that of (n, α) reaction but somewhat larger than that of $(\text{n}, \alpha 2\text{n})$ reaction. Therefore, the production rate of P^{33} may be expected to be ~ 0.1 atoms/min/kg Cl.

The production of P^{32} in the reaction $\text{S}^{32}(\text{n}, \text{p}) \text{P}^{32}$ can be calculated more accurately. The excitation function in the energy range 2 to 15 Mev is well established (BNL-325). The threshold is at ~ 2 Mev and the excitation curve is fairly flat at 350 mb in the energy range 5 to 15 Mev. The cross section at energies higher than 15 Mev is not known. But, in common with most (n, p) reactions, the cross section may be expected to drop to a low value at ~ 30 Mev; the value at 90 Mev is certainly small (< 40 mb; Knox, [1949]). Using this excitation function and the neutron data of Hess, Patterson, Wallace, and Chupp [1959] we estimate the production rate of P^{32} to be 1.0 atoms/min/kg S.

An estimate of production rate of Cl^{39} in the reaction $\text{Ar}^{40}(\text{n}, \text{pn}) \text{Cl}^{39}$ may be made by assuming that the excitation function for this reaction is similar to that for (p, pn) reactions in elements close to argon. The excitation functions for the (p, pn) reaction in different elements have similar shape but considerably

different value for the absolute cross section; the peak value at 20 to 30 Mev for Na^{23} , Co^{59} , Cu^{63} , and Cu^{65} being 120, 770, 570, 530 mb, respectively [Meadows and Holt, 1951a, b; Meadows, 1953; Wagner and Wiig, 1954; Coleman and Tewes, 1955]. We take 200 mb as the peak value for the reaction $\text{Ar}^{40}(\text{n}, \text{pn}) \text{Cl}^{39}$. This yields a production rate of ~ 0.2 atom of Cl^{39} /min/kg Ar. Production of Cl^{38} in the reactions $\text{Ar}^{40}(\text{n}, \text{p}2\text{n}) \text{Cl}^{38}$ and $\text{Ar}^{40}(\text{n}, 2\text{pn}) \text{S}^{38} \rightarrow \text{Cl}^{38}$ is expected to be somewhat smaller, since these reactions have a little higher threshold and lower cross section than those for (n, pn) reactions.

The production rate of Au^{198} and Au^{199} in $(\text{n}, \text{p}2\text{n})$, especially (n, pn) and $(\text{n}, \text{p}2\text{n})$, reactions in different isotopes of mercury may be obtained by using cross sections for equivalent $(\text{p}, \text{p}2\text{n})$ reactions in heavy elements (Cs^{133} , Ta^{181} , Bi^{209}) [Fink and Wiig, 1954; Markowitz, Rowland, and Friedlander, 1958; Bennet, 1954]. We take 200 to 400 mb as the average total cross section in the energy range 20 to 100 Mev. This gives the combined production rate for Au^{198} and Au^{199} as 0.05 to 0.1 atoms/min/kg Hg.

Negative-muon-induced activities. Cl^{39} : Winsberg [1956] calculated the production of Cl^{39} by the capture of cosmic-ray negative muons in atmospheric argon. Employing the same procedure, we estimate that 0.26 negative muons come to rest per minute in 1 kg of matter at sea level. The probability that a stopped negative muon will undergo nuclear capture in argon is close to 0.8. Therefore, the rate of nuclear capture in argon is 0.26×0.8 /min/kg. For calculating the production of Cl^{39} , we must also know the probability of emission of single neutron per nuclear capture. The individual probabilities of emission of 0, 1, 2, 3, and 4 neutrons from the capture of negative muons in iodine have been estimated by Winsberg [1954] to be about 8, 43, 34, 11, and 2 per cent, respectively. It is perhaps not unreasonable to take a value of 43 per cent for emission of a single neutron in the case of argon as well. The rate of production of Cl^{39} at sea level should therefore be $0.26 \times 0.8 \times 0.43$, i.e. 0.09 atoms/min/kg.

Au^{198} and Au^{199} : The probability of nuclear capture for stopping negative muons in mercury is 0.96. The rate of nuclear captures, therefore, should be 0.26×0.96 /min/kg Hg. We now require to know the percentage of nuclear

captures leading to the production of Au^{198} and Au^{199} . Mercury has seven stable isotopes; their natural abundances are Hg^{204} (6.8%), Hg^{200} (29.8%), Hg^{201} (13.2%), Hg^{202} (23.1%), Hg^{203} (16.9%), Hg^{198} (10.0%), and Hg^{196} (0.15%). Au^{198} will result from nuclear captures in Hg^{198} , Hg^{199} , Hg^{200} , Hg^{201} , and Hg^{202} provided that result in the emission of 0, 1, 2, 3, and 4 neutrons respectively. Using the probabilities of emission for 0, 1, 2, 3, and 4 neutrons as given by Winsberg, we estimate that only 18 per cent of nuclear captures in natural mercury leads to the production of Au^{198} . From a similar calculation, the fraction for Au^{199} turns out to be 12 per cent. The combined production for Au^{198} and Au^{199} should therefore be equal to $0.26 \times 0.96 \times 0.35$, i.e. 0.097 atoms/min/kg Hg. The detection of other isotopes of gold was not possible under our experimental conditions; therefore, no calculation is attempted.

The production of radiogold under 30 meters of rock should be comparatively very small. The flux of slow muons at this depth, as inferred from differential range spectrum of muons at sea level [Rossi, 1948] should be a factor of ~ 10 lower than that at sea level. The rate of radiogold production should therefore be ~ 0.002 atoms/min/kg Hg.

Acknowledgments. We are indebted to Professor James R. Arnold for very valuable suggestions during the course of this work and to Professor Edward D. Goldberg for stimulating discussions. Skillful assistance of Mr. M. A. Anderson and Mr. J. Lucas is gratefully acknowledged. We are thankful to Mrs. Roger Revelle and Mrs. H. T. Byington for extending the facilities for irradiation of mercury targets in the swimming pool and in the cave, respectively.

This work was supported by a contract with the U. S. Air Force Geophysics Research Directorate.

REFERENCES

- Bennet, W. E., Spallation studies of bismuth, *Phys. Rev.*, **94**, 997-999, 1954.
- Coleman, G. H., and H. A. Tewes, Nuclear reactions of Cu with various high energy particles, *Phys. Rev.*, **99**, 288-289, 1955.
- Fink, R. W., and E. O. Wiig, Reactions of cesium with protons at 60, 80, 100, 150, and 240 Mev, *Phys. Rev.*, **96**, 185-187, 1954.
- Harding, J. B., S. Lattimore, T. T. Li, and D. I. Perkins, Absorption of star-producing radiation under ice, *Nature*, **163**, 319-320, 1949.
- Hess, W. N., H. W. Patterson, R. Wallace, and E. L. Chupp, Cosmic ray neutron energy spectrum, *Phys. Rev.*, **116**, 445-457, 1959.

- M., J. P. Shedlovsky, and J. R. Arnold, Active species produced by cosmic rays in meteorites, *Geochim. et Cosmochim. Acta*, **25**, 154, 1961.
- W., Spallation yields from chlorine with 1 Mev protons, *NYO-6627*, Carnegie Institute of Technology, Pittsburgh, Pa., 1961.
- J., Relative cross sections for nuclear reactions induced by high energy neutrons in elements, *Phys. Rev.*, **75**, 537-541, 1949.
- J. R. Arnold, and M. Honda, Cosmic ray reaction rates of Be^7 in oxygen, and P^{32} in argon at mountain altitudes, *Phys. Rev.*, **118**, 1626-1632, 1960.
- W., S. S., F. S. Rowland, and G. Friedlander, (p, pn) reactions at proton energies from 0.5 to 8.0 BeV, *Phys. Rev.*, **112**, 1295-1302, 1958.
- L., Spallation of Cu with high energy protons, *Phys. Rev.*, **88**, 225-227, 1952.
- J. W., Excitation functions for proton-reactions with copper, *Phys. Rev.*, **91**, 1000-1003, 1953.
- J. W., and R. B. Holt, Some excitation functions for protons on magnesium, *Phys. Rev.*, **83**, 1257, 1951a.
- J. W., and R. B. Holt, Excitation functions for proton reactions with sodium and magnesium, *Phys. Rev.*, **83**, 47-49, 1951b.
- Paul, E. B., and R. L. Clarke, Cross section measurements of reactions induced by neutrons of 14.5 MeV energy, *Can. J. Phys.*, **31**, 267-277, 1952.
- Rossi, B., Interpretation of cosmic ray phenomena, *Revs. Modern Phys.*, **20**, 537-583, 1948.
- Rossi, B., High energy particles, *Prentice-Hall Physics Series*, Prentice-Hall, New York, 441 pp., 1952.
- Sens, J. C., Capture of negative muons by nuclei, *Phys. Rev.*, **113**, 679-687, 1959.
- Tennet, R. M., The absorption and decay of negative muons, *Progress in Elementary Particles and Cosmic Ray Physics*, **5**, 364-408, North Holland Publishing Co., Amsterdam, 1960.
- Wagner, G. D., and E. O. Wiig, Reactions of cobalt with protons at 60, 100, 170, and 240 MeV, *Phys. Rev.*, **96**, 1100-1103, 1954.
- Winsberg, L., Interactions of negative muons with iodine, *Phys. Rev.*, **95**, 205-209, 1954.
- Winsberg, L., The production of chlorine-39 in the lower atmosphere by cosmic radiation, *Geochim. et Cosmochim. Acta*, **20**, 183-189, 1956.

(Manuscript received May 16, 1961.)

Radioactive Species Produced by Cosmic Rays in Bruderheim and Other Stone Meteorites

M. HONDA, S. UMEMOTO, AND J. R. ARNOLD

*School of Science and Engineering, University of California, San Diego
La Jolla, California*

Abstract. In this paper we report measurements of the activity of 16 radioactive nuclides in the chondritic meteorite Bruderheim, which fell on March 4, 1960. The data are compared with those for the iron Aroos and other stone, iron, and stony-iron meteorites. For nuclides of $A \geq 46$, the activities normalized to iron and nickel content are closely similar. For lower A , the activities in stone are mainly produced in lighter elements. The spectrum of bombarding particles is shown to be closely similar in irons and stones. The results are consistent with a constant cosmic-ray intensity. Cosmic-ray bombardment ages of 30×10^6 years for Bruderheim and 130×10^6 years for Admire are calculated from rare-gas data and the activities of Na^{22} and Cl^{36} .

Record of cosmic-ray bombardment in meteorites is in the form of rare stable and radioactive nuclei produced by transmutation. Measurements of the concentration of these nuclei have been made, and they have been used to draw conclusions about both cosmic rays and meteorites. The field has recently been reviewed by Arnold [1961].

Shedlovsky, and Arnold [1961] have measured the activity of several long-lived radioactive nuclides in four iron meteorites, and Arnold [1961] have measured a large number of short- and long-lived species in the stony iron Aroos. From these data and other workers it has been concluded [Honda, and Lal, 1961] that the cosmic-ray intensity, averaged over the half-life of each nuclide, has been constant within the limits of theory and experiment. It has been possible to account with adequate precision for the production rates and concentrations of radioactive and stable species in iron meteorites. Predictions of this model have been verified by Pfeffer and Honda, 1961].

It is also of interest, as part of the same program, to measure the concentrations of as many radioactive nuclides as possible in stony meteorites [Honda and Kohman, 1958]. This paper is concerned with the measurement of a number of radioactive species in the recently fallen chondrite Bruderheim. We also report our measurements of long-lived activities in the chondrite Achilles, and in the separated

stone and iron fractions of the pallasites Admire and Brenham. A description of these meteorites is given in the Appendix. We will discuss the meaning of these results for the energy distribution of the bombarding particles, the constancy of the cosmic-ray intensity in time and space, and the cosmic-ray bombardment ages of these meteorites.

The Bruderheim chondrite fell on March 4, 1960, near Edmonton, Alberta, Canada. The fall consisted of many stones; the total weight exceeded 100 kg. This is a typical gray chondrite whose composition is very close to the mean of the low-iron group of Urey and Craig [1953] [Baadsgaard, Campbell, Cumming, and Folinsbee, 1961; Duke, Maynes, and Brown, 1961]. Through the courtesy of Dr. Folinsbee, we received a large specimen on May 9, 1960.

Using this sample we have measured the content of 16 radioactive nuclides, including most of those reported for Aroos and Fe^{55} (an upper limit was also set for Be^7). Wet chemical procedures were used throughout. Isolation steps and counting methods were similar to those for Aroos. The activities of species of $A \geq 46$, for a given weight of Fe, Co, or Ni, are close to those found in Aroos, except for special cases such as the neutron capture species Co^{60} . The measurements of long-lived Cl^{36} , Al^{26} , and Be^{10} agree with those for the other stones, when chemical composition is taken into account.

These data, together with the rare-gas concentrations, yield bombardment ages of $30 \times$

TABLE 1. Content of Cosmic-Ray-Produced Radioactivities in Bruderheim*

Nuclide	$t_{1/2}$	dpm/kg at Time of Fall
Be ⁷	53 d	≤100
Be ¹⁰	2.5×10^6 y	19 ± 2
Na ²²	2.58 y	90 ± 10
Al ²⁶	7.4×10^5 y	60 ± 6
Cl ³⁶	3.1×10^5 y	7.5 ± 0.8
Sc ⁴⁶	83.8 d	6.2 ± 0.6
Ti ⁴⁴	~200 y	2.0 ± 0.2
V ⁴⁸	16.0 d	34 ± 7
V ⁴⁹	330 d	34 ± 5
Cr ⁵¹	27.8 d	110 ± 27
Mn ⁵³	$\geq 2 \times 10^6$ y	85 ± 17
Mn ⁵⁴	308 d	100 ± 13
Fe ⁵⁵	2.6 y	340 ± 80
Co ⁵⁶⁺⁵⁸	~74 d	14 ± 4
Co ⁵⁷	240 d	11 ± 1
Co ⁶⁰	5.26 y	9 ± 1
Ni ⁵⁹	8×10^4 y	12 ± 3

* Other data: Ar³⁹ 10 ± 1 ; Ar³⁷ 23 ± 4 ; T 260 ± 30 [Fireman and DeFelice, 1961].

10^6 years for Bruderheim and 130×10^6 years for Admire.

EXPERIMENTAL PROCEDURES

Chemical treatment. A total of 850 grams of Bruderheim was dissolved in HF and H₂SO₄. After drying, the cake was extracted with water. The residue was fused with NaOH and extracted with HCl; Cr and other elements were recovered from this fraction (and later combined with the proper fractions from the main solution). Carriers of Sc, V, and Be were added to the main solution. Fe was separated by ether extraction. A large precipitate of Mg and Al was removed from saturated HCl solution. We note that throughout the process the removal of Mg without excessive losses of other elements was a major difficulty. Hydroxides were precipitated with NH₄OH, leaving most Ni in solution. A second MgCl₂ precipitation was carried out from saturated HCl. It was necessary later to treat the chloride precipitates to recover various hydroxide group elements. The solution was neutralized again with NH₄OH, and the precipitate was dissolved in 1 M HCl and H₂O₂. The cations were separated on a 1-liter cation-exchange column, eluting with 1-6 M HCl. The order of recovery was V, Ti, Be, Mn, Al, Cr, and Sc. A sulfide precipitation

removed Ni and Co from the NH₄OH solution. Na was recovered from the supernate by exchange, after treatment with Ba(OH)₂.

For Cl a separate sample of 80 grams was fused with NaOH. After extraction the residue was attacked with dilute HNO₃. Chloride carrier was added to the combined solutions, and AgCl precipitated.

The Achilles meteorite (100 grams) was attacked with HNO₃ and HF at room temperature. The Cl fraction was recovered from the filtrate as AgCl. The residue from the treatment was combined with the filtrate and treated further at high temperature with H₂O. The water extract was saturated with HCl at low temperature, and a white precipitate containing Al was separated. Fe^{III} was removed by anion exchange. After addition of excess NaOH and Be carrier, Be and Al remained in the filtrate.

The stone phase of Admire (75 grams) was treated with HF and H₂SO₄. Fe^{III} was removed by ether extraction in HCl solution. Al was precipitated using HCl-ether. Be was recovered from the filtrate.

The metal phases of Admire (340 grams) and Brenham (200 grams) meteorites were treated by a procedure similar to those reported previously [Honda, Shedlovsky, and Arnold, 1961]. They were dissolved in nitric acid, and Cl³⁶, Fe⁵⁵, Al²⁶, K⁴⁰, Mn⁵³, and Ni⁵⁹ were separated.

Counting methods. Our counting methods for β , γ , and X radiation, are described in earlier papers [Honda and Arnold, 1961]. The counting samples were usually more massive than those from Aroos. The self-absorption factor is especially important for Na²², Al²⁶, Cr⁵¹, and Co⁶⁰. For Fe⁵⁵ and Ni⁵⁹ the samples were never 'infinitely thick.' For Na²² and Al²⁶ in Bruderheim

TABLE 2. Content of Cosmic-Ray-Produced Radioactivity in Some Stone and Stony-Iron Meteorites, dpm/kg

	Achilles	Admire Stone	Admire Metal	Brenham Metal
Be ¹⁰		14 ± 2	1.7 ± 0.3	0.2 ± 0.1
Al ²⁶	50 ± 5	43 ± 4	1.5 ± 0.5	0.1 ± 0.1
Cl ³⁶	6.0 ± 0.6		7.4 ± 0.9	0.1 ± 0.1
K ⁴⁰			1.1 ± 0.4	
Mn ⁵³			200 ± 20	<15
Ni ⁵⁹			300 ± 30	<20

isotrons were counted in both β and γ

nuclides were identified by (1) recycling
ant specific activity; (2) absorption
hents for Be^{10} , Na^{22} , Sc^{46} , and Co^* ;
measurements for Sc^{46} , V^{48} , V^{49} , Cr^{51} ,
and Co^* .

RESULTS AND DISCUSSION

Results for Bruderheim are shown in Ta-
ble 2, and for Aroos, in Table 3. The accuracy of
the data is calculated from the standard devia-
tion statistics. Ten per cent is taken
as the maximum, including other sources of error
(Honda and Arnold, 1961).

Because of the shorter time between fall and
analysis as well as the greater size of the
specimen, the V^{48} and Cr^{51} data for Bruderheim
are more reliable than those for Aroos. The figure for
 Be^{10} is only an upper limit. The chemical
separation of each element, except Sc, V, and Be, is
based on the analytical data. We have as-
sumed concentrations of 10 ppm Sc, 30 ppm V,
and negligible Be, in addition to added carrier.

Table 2 shows the results obtained in other
specimens. Within experimental error no activ-
ities were present in our specimen of Brenham metal.

(Private communication) reports the ab-
sence of cosmogenic rare gases. Since helium
was measured in another specimen of this
type, we must conclude that our sample
was not deeply within the original mass. The
data for Admire metal are about half those
for Bruderheim, except for the high value of Ni^{59} . The
data for Achilles and Admire stone compare
favorably with those in Bruderheim.

The Al^{26} measurements are consistent with
those reported by Ehmman and Kohman
for the Al^{26} content of chondrites has also
been checked by other workers using different
methods [Van Dilla, Arnold, and Anderson,
1961]. For Be^{10} , however, much lower values
were reported by Ehmman and Kohman: $1.6 \pm$

0.5 ± 0.5 dpm/kg in Plainview and
Aroos, respectively. Their values are com-
parable to those obtained for small iron meteor-
ites because of the high abundance of oxygen,
which increases the production rate in chondrites must be sub-
stantially higher. Their low figures might possi-
bly be due to incomplete separation of Be from

Vilcsek and Wänke [1960] reported the Na^{22}
content of the chondrite Breitscheid (fell in
1956). Their figure, 89 ± 15 dpm/kg, is in good
agreement with ours. The same authors [1961]
tried to measure Cl^{36} in Bruderheim by extrac-
tion with dilute acid from a powdered sample.
They obtained 5.7 ± 0.4 dpm/kg, which is close
to our value even before considering possible
incompleteness of the extraction.

Rowe and van Dilla [1961] report values of
60 dpm/kg Al^{26} , 90 dpm/kg Na^{22} , and 82 dpm/kg
 Mn^{54} in Bruderheim. These values, obtained by
 γ -ray spectrometry on an intact specimen, are
in excellent agreement.

Comparison with Aroos and other iron meteor-
ites. The activities measured in Bruderheim can
be compared with those of a small iron meteorite
such as Aroos. In Table 3 we have multiplied
the activities found in Aroos [Honda and Arnold,
1961] by a factor of 0.24 (0.18 for Ni and Co
nuclides). The factor 0.24 corresponds to the
sum of the Fe, Co, and Ni concentrations in
Bruderheim in weight per cent, whereas 0.18 is

TABLE 3. Content of Radioactivity in
Bruderheim and Aroos Meteorites at
Time of Fall, dpm/kg

Nuclide	Bruder- heim	0.24 \times Aroos	Ratio
Be^{10}	19	1	19 ± 2.7
Na^{22}	90	0.4	200 ± 30
Al^{26}	60	0.9	70 ± 12
Cl^{36}	7.5	3.3	2.3 ± 0.3
Ar^{37}	23	5*	4.6 ± 0.6
Ar^{39}	10	4*	2.5 ± 0.4
Sc^{46}	6.2	7.2	0.86 ± 0.12
Ti^{44}	2.0	1.1	1.8 ± 0.3
V^{48}	34	22	1.5 ± 0.7
V^{49}	34	38	0.89 ± 0.16
Cr^{51}	110	65	1.7 ± 0.8
Mn^{53}	85	124	0.69 ± 0.10
Mn^{54}	100	113	0.89 ± 0.12
Fe^{55}	340	380†	0.89 ± 0.45
		0.18 \times Aroos	
Co^{56+58}	14	22	0.64 ± 0.20
Co^{57}	11	16	0.69 ± 0.10
Co^{60}	9	3	3 ± 0.4
Ni^{59}	12	11	1.1 ± 0.3

* From Fireman and DeFelice [1960].

† The value of 1600 ± 600 dpm/kg is due to M.
Honda (unpublished).

TABLE 4. Comparison of Specific Activities in Stone and Metal

dpm/kg Target	Stone			Metal		
	Bruderheim	Achilles	Admire Stone	dpm/kg Target	Admire Metal	Aro.
Be ¹⁰ /O	53 ± 5		32 ± 4	Be ¹⁰ /Fe + Ni	1.7 ± 0.3	4.1 ±
Al ²⁶ /Si + Al	300 ± 30	250 ± 25	220 ± 22	Al ²⁶ /Fe + Ni	1.5 ± 0.5	3.6 ±
Cl ³⁶ /Ca + K*	320 ± 64	210 ± 40		Cl ³⁶ /Fe + Ni	7.4 ± 0.9	14 ±
Mn ⁵³ /Fe + Ni	420 ± 100			Mn ⁵³ /Fe + Ni	200 ± 20	515 ±
Relative bombardment intensity	0.82	0.65	0.55		0.45	1.00

* Corrected for contribution from target Fe + Ni.

the relative content of Co and Ni. The ratios are not far from unity between Sc⁴⁶ and Ni⁶⁰ except for Co⁶⁰ and Ti⁴⁴. The weighted average (excluding these two nuclides) is 0.82. There is no trend with ΔA (defined as $A_{\text{target}} - A_{\text{product}}$). We conclude that the relative spectrum of nuclear-active particles [Arnold, Honda, and Lal, 1961] in small iron meteorites and in Bruderheim is effectively the same for the production of species from Sc⁴⁶ to Ni⁶⁰ and that the flux in our Bruderheim sample was about 80 per cent of that in our sample of Aroos. The excess of Co⁶⁰ we attribute to a higher slow neutron flux in Bruderheim. This is to be expected, since neutrons are more effectively moderated by low-Z elements. The activity of Co⁶⁰ in Bruderheim is about 10⁴ dpm/kg Co. The small excess of Ti⁴⁴ may be ascribed to production from Ti and Cr.

For the other nuclides in this group, neutron capture and production from higher elements are not expected to be important.

From Be¹⁰ to Ar³⁹, the large differences observed must be attributed to production from other major constituents closer to the product nuclides.

Comparison of production in iron and stone phases. In Table 4 we compare the activities of some long-lived isotopes in Bruderheim, Achilles, the two phases of Admire, and Aroos. The activities have been divided by the weight per cent of the target elements which are most important for each isotope. This corrects (accurately enough for present purposes) for the effect of changes of composition. The last row gives the relative bombardment intensities in the various samples as determined from activity

ratios. The value of 0.82 for our sample of Bruderheim is taken from the previous section.

The most important conclusion from the data and those in Table 3 is that the spectrum of bombarding particles in Bruderheim or Admire is closely similar to that in Aroos, a typical small iron meteorite. One direct demonstration of this is the lack of a trend in the ratios of the observed activities in Bruderheim and Aroos in going from Sc⁴⁶ to Fe⁵⁵. The former was produced mainly at energies of several hundred Mev, the latter mainly in the region of 10 to 20 Mev. A relative change of 40 per cent in the two differential spectra over the range from 10 to 500 Mev [Arnold, Honda, and Lal, 1961] would have produced a visible trend.

The comparison of the bombardment intensities in the two phases of Admire allows us to extend this result to higher energies. The species Be¹⁰ and Al²⁶ are produced in iron mainly in the beV region. In stone they are produced mainly in O and Si respectively, by reactions that have large cross sections at low energies. They are therefore, like Mn⁵⁴ in iron, produced mainly by low-energy particles. The relative bombardment intensities of 0.55 and 0.45 in the two phases are nearly the same. If the difference is significant, it can be accounted for by a difference of about 20 per cent in the spectra in our specimens of Admire and Aroos between low energies at the beV region. The difference would be in the direction of a 'softer' spectrum in Admire. The lower bombardment intensity in our specimen of Admire is consistent with a small effect in this direction.

Bruderheim and Achilles are low-iron chondrites. We may expect the relative production

5(a). Cosmic-Ray-Produced Rare-Gas
Content: Unit: 10^{-8} cc/g (NTP)

der- m*	Admire Stone*	Admire Metal†	Bren- ham Stone*	Brenham Metal†
22‡	0.49	3.9	0	≤ 0.03
80‡	0.34	2.5	0	
6	41	0.76		≤ 0.01
5	101	63	0.2	≤ 0.5

TABLE 5(b). Bombardment Age:
Unit: 10^6 years

Bruderheim	Admire Stone	Admire Metal
33 \pm 8		140 \pm 20
27 \pm 4		
	120 \pm 20	
30 \pm 4		130 \pm 14

Stauffer (private communication). Estimates: ± 10 –15 per cent for Ne and He, ± 20 for Ar.

and Nier [1961].

data on the Ar isotopes have been corrected for diatomic gas, using the procedure described by [1961]. The uncorrected values were Ar^{38} , 1.39.

All species to be similar in other representatives of this class. The production in stones and phases of substantially different composition may be estimated from the data in Table 4. We note that an appreciable fraction of the Ar isotopes is produced from the reaction of production rates. In another study [Arnold, Honda, and Lal, 1961] we have used a method of calculating production rates for various species, using cross sections along with experimental data of cosmic-ray physics. The resulting relative and absolute production rates are in good agreement with the data for other iron meteorites. Since it has been demonstrated above that the relative differential spectra of the bombarding particles in the samples of Bruderheim and Admire were closely similar to those of the Aros, the same method is also applicable to the Aros. However, the necessary cross-section data for the targets O, Mg, Si, Ca, and Fe are not yet available, especially at the energies that are most important.

The consistency of the production rates may be seen from the data on Bruderheim in Table 4. The production rates of Ar^{37} and Ar^{39} are discussed by Stoenner, Schaeffer, and Davis [1960] and Fireman and DeFelice [1961]. The production rates of Al^{28} and Cl^{36} , normalized to the weight of the main target elements, are similar to those of Mn^{54} and Mn^{55} , species of similar ΔA produced in Fe. The ratio $\text{Na}^{23}/\text{Mg} + \text{Na}$ amounts to 570 dpm/kg; since a significant contribution from Si and Al must be present this value is also close to the others. The value of Be^{10}/O is much lower than that for the comparable case of V^{50}/Fe , as is to be expected, since the cross section for Be^{10} production by 220-Mev protons in a CNO target is quite low, in fact several times lower than that for Be^7 , according to unpublished data of Honda and Lal. Thus the production rates of the species produced in light elements appear to be consistent with the others.

The data on Aros and other iron meteorites led us to the conclusion that the cosmic-ray intensity, averaged over the half-life of each radioactive product nuclide, has been nearly constant. The data on Bruderheim and other stones also appear, from this discussion, to be consistent with the conclusion that secular equilibrium prevails.

Cosmic-ray-produced isotopes of heavier elements. The production of Fe^{55} by the $(n, 2n)$ reaction reaches almost 2000 dpm/kg, or 2 dpm/g target. The product of a low-energy reaction such as (n, p) , $(n, 2n)$, or (n, pn) on another element might possibly be made at a rate as high as 10 atoms/g min. The activity of the (n, γ) product Co^{60} reaches 10 atoms/g min in Bruderheim; still higher values are possible, of course, if the capture cross section is very large. Over a bombardment period of 10^6 years, a rate of 10 atoms/g min would result in the transmutation of about 4×10^{-8} of the target atoms. This might in special cases give a detectable result even for trace-element targets.

Rare-gas content and cosmic-ray age. A comparison of the activity data with the concentrations of stable rare gases permits the calculation of cosmic-ray ages. Rare-gas measurements on our samples were made by P. Signer and H. Stauffer (unpublished). The data are given in

Table 5(a). The calculated bombardment ages are shown in Table 5(b).

For Bruderheim, we have used the pairs Na^{22} - Ne^{22} and Cl^{36} - Ar^{36} . We have assumed 1:1 for the direct production ratio Na^{22} : Ne^{22} . For Cl^{36} : Ar^{36} , we have assumed 1:1 for production from $\text{Ca} + \text{K}$, and the usual value of 4:1 for production from Fe. From these we calculate a ratio of 6:4. For Cl^{36} : Ar^{36} in Admire metal, the usual ratio of 4:1 has been taken. The Al^{26} - Ne^{22} age in the stone phase has been calculated as follows. We estimate that the ratio Al^{26} : Na^{22} is one-third lower in the Admire stone phase than in Bruderheim, because the Si:Mg ratio in Admire stone is only about half as great. On this basis the Na^{22} content of Admire stone at the time of fall was the same as that of Bruderheim, and the age therefore four times greater.

The values in Table 5 are mutually consistent. Incorrect assumptions about production ratios, and possible diffusion losses, are possible sources of error. Vilček and Wänke [1960] have obtained an age of 30×10^6 years for the chondrite Breitscheid, using the same assumptions for the pair Na^{22} - Ne^{22} . Values around 20×10^6 years or less have been obtained for other chondrites by the H^3 : He^3 method [Geiss, Hirt, and Oeschger, 1960]. The data are not sufficient to permit a useful discussion of this difference.

APPENDIX. METEORITE SAMPLES

Admire. Stony-iron meteorite, Brecciated pallasite. Lyon County, Kansas, USA. Found 1881. Total mass collected more than 50 kg. Our specimen was a composite of Ward Catalog no. S 1330 (271 g); S 1327 (337 g); S 1335 (312 g): Measured density 4.55 (S 1335).

Brenham. Stony-iron meteorite, pallasite. Brenham Township, Kiowa County, Kansas, USA. Found 1882. Total mass collected several tons. Our specimen: Ward Catalog no. 10:50 (365 g); no. 10:54 (259 g). Measured density 4.88-4.98 (corresponding to a 1:1 weight ratio of stone and metal phases).

Achilles. Stone meteorite, veined crystalline chondrite. Rawlins County, Kansas, USA. Found 1924. Stone. Ward Catalog no. S 1826 (445 g).

Acknowledgments. We are indebted most of all

to R. E. Folinsbee for making the sample of Bruderheim available to us. We are also indebted to him and to Harrison Brown for analytical data on this meteorite in advance of publication, to Signer and H. Stauffer for the rare-gas measurements, and to Maurice Anderson for assistance in the chemical work. This research was carried out under a grant from the National Aeronautics and Space Administration.

REFERENCES

- Arnold, J. R., *Ann. Rev. Nuclear Sci.*, **11**, 1961.
 Arnold, J. R., M. Honda, and D. Lal, Recording cosmic-ray intensity in the meteorites, *J. Geophys. Research*, **66**(10), 1961.
 Baadsgaard, H., F. A. Campbell, G. L. Cummins, and R. E. Folinsbee, The Bruderheim meteorite, *J. Geophys. Research*, **66**(10), 1961.
 Duke, Michael, Donald Maynes, and Harrison Brown, The petrography and chemical composition of the Bruderheim meteorite, *J. Geophys. Research*, **66**(10), 1961.
 Ehmann, W. D., and T. P. Kohman, *Geochim. et Cosmochim. Acta*, **14**, 340, 364, 1958.
 Fireman, E. L., and J. DeFelice, Argon 37, argon 39, and tritium in meteorites and the spatial constancy of cosmic rays, *J. Geophys. Research*, **65**, 3035-3041, 1960.
 Fireman, E. L., and J. DeFelice, Tritium, argon 37, and argon 39 in the Bruderheim meteorite, *J. Geophys. Research*, **66**(10), 1961.
 Geiss, J., B. Hirt, and H. Oeschger, *Helv. Phys. Acta*, **33**, 590, 1960.
 Honda, M., and J. R. Arnold, *Geochim. et Cosmochim. Acta*, **23**, 219, 1961.
 Honda, M., J. P. Shedlovsky, and J. R. Arnold, *Geochim. et Cosmochim. Acta*, **22**, 133, 1961.
 Rowe, M., and M. A. van Dilla, On the radioactivity of the Bruderheim chondrite, *J. Geophys. Research*, **66**(10), 1961.
 Signer, P., and A. O. Nier, The measurement and interpretation of rare gas concentrations in iron meteorites, preprint, 1961.
 Stauffer, H., Primordial argon and neon in carbonaceous chondrites and ureilites, *Geochim. et Cosmochim. Acta*, 1961.
 Stauffer, H., and M. Honda, Cosmic-ray production of V^{50} and K^{40} in the iron meteorite Aroos, *J. Geophys. Research*, **66**, 3584-3586, 1961.
 Stoenner, R. W., O. A. Schaeffer, and R. Davy Jr., Meteorites as space probes for testing the spatial constancy of cosmic radiation, *J. Geophys. Research*, **65**, 3025-3034, 1960.
 Urey, H. C., and H. Craig, *Geochim. et Cosmochim. Acta*, **4**, 36, 1953.
 Van Dilla, M. A., J. R. Arnold, and E. C. Anderson, *Geochim. et Cosmochim. Acta*, **20**, 115, 1960.
 Vilček, E., and H. Wänke, *Z. Naturforsch.*, **15a**, 1004, 1960; **16a**, 379, 1961.

(Manuscript received June 26, 1961;
 revised July 26, 1961.)

Tritium, Argon 37, and Argon 39 in the Bruderheim Meteorite

E. L. FIREMAN AND J. DEFELICE

*Smithsonian Astrophysical Observatory
Cambridge, Massachusetts*

Abstract. Tritium and argon 39 were measured in whole rock samples and in separated phases of the Bruderheim meteorite. Argon 37 was measured in a Bruderheim whole rock sample. There were 260 ± 30 tritium decays/kg min, 10 ± 1 argon 39 decays/kg min, and 10 ± 4 argon 37 decays/kg min in the whole rock samples. The tritium/argon 39 ratio was 26 ± 5 , and the argon 37/argon 39 ratio was 2.2 ± 0.4 ; these ratios are identical to those measured in the Hamlet meteorite. In the nonmagnetic (silicate) phase, there were 290 ± 30 tritium decays/kg min and 4.8 ± 0.6 argon 39 decays/kg min. The tritium/argon 39 ratio of 60 ± 10 in this phase agrees with that expected from the chemical composition and the high-energy neutron cross sections. In the magnetic (metallic) phase, there were 90 ± 20 tritium decays/kg min and 36 ± 6 argon 39 decays/kg min. The high argon 39 content of the metallic phase proves that the argon 39 is produced chiefly from the element iron and that chondritic meteorites are in general less shielded from cosmic rays than metallic meteorites. The tritium/argon 39 ratio of 2.5 ± 0.5 in the metallic phase shows that much of the tritium is lost in this phase; however, since less than 10 per cent of the meteorite is in the metallic phase, tritium loss from the metal does not introduce a serious error in the tritium-helium 3 exposure age. An exposure age of 40 ± 10 million years was obtained by combining our tritium value for the whole rock sample with a helium 3 value of 50×10^{-8} cc/g. An exposure age of 4 ± 4 m.y. was obtained by combining our argon 39 value with an argon 38 value of 1.4×10^{-8} cc/g.

Introduction. Tritium (12.4-year half-life) was measured in about 20 meteorites [Fireman and Schwarzer, 1957; Begemann, Geiss, and Fireman, 1957; Begemann, Eberhardt, and Hess, 1957; Bebel and Schmidlin, 1960; Fireman and DeFelice, 1960b]; argon 39 (325-year half-life) was measured in about 10 meteorites [Fireman and DeFelice, 1958; Sprengel, Davis, and Wügg, 1959; Fireman and Vilcsek, 1959; Fireman and DeFelice, 1960; Stoenner, Schaeffer, and Davis, 1960]; argon 37 (34-day half-life) has been measured in only two meteorites, Hamlet and Aroos [Fireman and DeFelice, 1960a; Stoenner, Schaeffer, and Davis, 1960]. The tritium content in stony meteorites is much lower than in metallic meteorites [Fireman and DeFelice, 1960b]. Stony and metallic meteorites have similar argon 37/argon 39 ratios. The argon 37 activity is twice the argon 39 activity in Hamlet (chondrite) and is larger than the argon 39 activity in Aroos (metal).

These results suggested that stony meteorites are less shielded from cosmic rays than metallic meteorites and that the energetic cosmic rays produce radioactive isotopes in meteorites

are constant to within a factor of 3. A few measurements, however, do not fit into this simple picture. The most striking is the very low tritium content of some metallic meteorites. Our measurements on the Bruderheim meteorite, which fell on March 4, 1960, greatly strengthen the conclusion that stony meteorites are less shielded from cosmic rays than metallic meteorites. The ratio of argon 37 to argon 39 in the Bruderheim meteorite is identical to that measured in the Hamlet meteorite. Since the calcium content is approximately the same in the two meteorites, the argon 37 from calcium may account for the excess argon 37 in both. It was not possible to measure the argon 37 in the separated phases and directly determine the argon 37 contribution from the calcium in the meteorite.

Experimental procedure. We have described the apparatus for the extraction of tritium and argon isotopes in earlier publications [Fireman and Schwarzer, 1957; Fireman and DeFelice, 1960b]. Our procedure, somewhat modified from our earlier methods, was as follows: Samples up to 20 grams are contained in a covered molybdenum crucible which is heated

TABLE 1. Tritium, Argon 37, and Argon 39 in the Bruderheim Chondritic Meteorite

Material	Weight, grams	Decays/kg min			Ratios of Radioactivity	
		Tritium	Ar ³⁷ *	Ar ³⁹	H ³ /Ar ³⁹	Ar ³⁷ /Ar ³⁹
Bulk	50	...	23 ± 4	10.5 ± 2	...	2.2 ± 0.2
Bulk	40	260 ± 30	...	10.0 ± 1.0	26 ± 5	
Magnetic phase	7.5	90 ± 20	...	36 ± 6	2.5 ± 0.5	
Nonmagnetic phase	39	290 ± 30	...	4.8 ± 0.6	60 ± 10	

* At time of fall.

by a radio-frequency heater. After evacuation, measured amounts of argon (0.2 sec) and hydrogen (2.0 and 3.0 sec) are added as carriers. The sample is melted, and the molten sample remains in contact with the carrier and evolved gas for 1 hour. The gas is then removed by means of a Toepler pump. The removal of the gas requires several hours; the sample is molten throughout this period. The hydrogen is diffused through a palladium thimble at approximately 800°C and collected. The amount of recovered hydrogen is always larger than the amount of hydrogen carrier, so that the yield cannot be directly determined in this step. The yield of the hydrogen extraction was later determined to be 90 per cent by melting a sample with tritiated hydrogen carrier. An additional amount of hydrogen carrier (1.0 to 2.0 sec) is added to the remaining gas, which is then passed over CuO at approximately 600°C to convert the hydrogen to water. The water is reconverted to hydrogen over a magnesium-mercury amalgam; this hydrogen is diffused through a hot palladium thimble and recovered. The hydrogen is quantitatively recovered in the CuO-Mg step. This hydrogen is either added to the previous hydrogen or counted separately in an 80-cc-volume G-M counter using argon-butene as a quench mixture. The efficiency of this counter for tritium was determined to be 80 per cent. The argon is removed and purified by putting the remaining gas in contact with zirconium at approximately 1000°C, absorbing the remainder on charcoal at liquid-nitrogen temperature, removing the argon at dry-ice temperature, and purifying it once more over hot zirconium. The argon yield was determined by comparing the final volume with the volume of carrier argon. The argon yield was normally larger than 90 per cent. This argon is added to a small G-M counter previously de-

scribed, using argon-butene as a quench mixture.

The counting was done with the counter in a low-level system removed from the extraction line. The low-level system consisted of 1 cc of mercury surrounded by an anti-coincidence ring inside a 6-inch iron shield. The background of the 80-cc hydrogen counter was 0.80 counts/min; the backgrounds of the three argon counters that were used were 50, 60, and 80 counts/min.

Experimental results. Our experimental results on Bruderheim are summarized in Table 1. We first measured a whole rock sample weighing 50 grams, whose argon activity was counted on April 26, 1960, 52 days after the date of fall. Figure 1 gives the argon activity from this sample as a function of time. The counting was continued until the middle of August, when only the argon 39 remained. When we subtracted the argon 39 activity, we find that the argon 37 decays with approximately the correct half-life. The ratio of argon 37 to argon 39 activity at the date of fall is 2.2 ± 0.2 by the method of least squares; the probable error is doubled in Table 1, which gives the value 2.2 ± 0.4 . The argon activity in a second whole rock sample was measured too late to obtain a reliable argon 37 value; the argon 39 in this sample was 10.0 decays/kg min, and the tritium activity was 260 ± 30 decays/kg min. To obtain absolute tritium activity in Bruderheim we use new correction factors for the chemical yield and counter efficiency. The chemical yield was determined to be 90 per cent instead of 100 per cent. The hydrogen counter efficiency based on tritium standard was 80 per cent instead of the previously calculated 95 per cent. These correction factors raise our previous values [Fireman and DeFelice, 1960a; 1960b] for tritium in chondrites by 30 per cent. Calibration of the e-

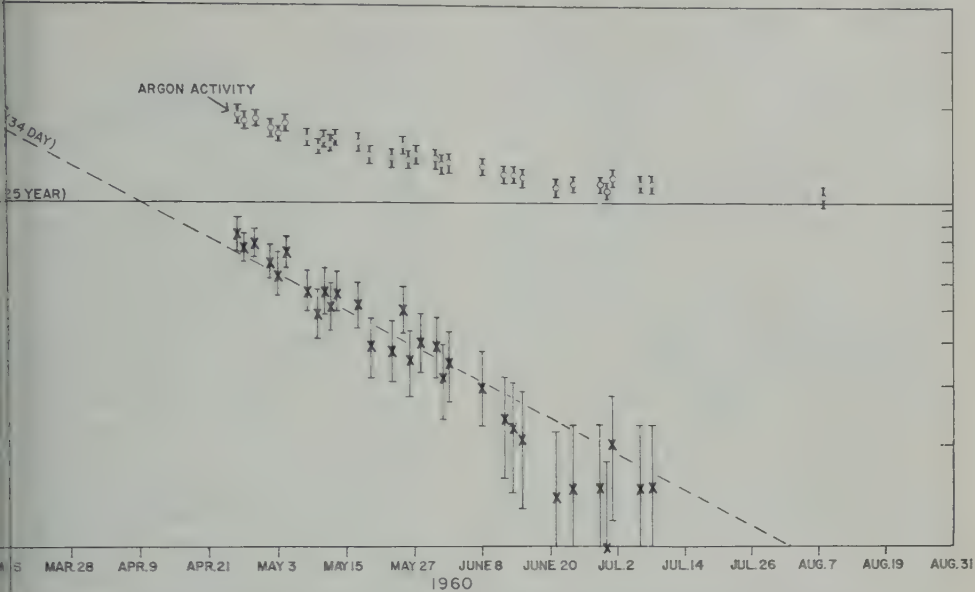


Fig. 1. Argon activity from the Bruderheim meteorite fall of March 4, 1960.

the two argon counters principally used
work gave values of 85 and 70 per cent.
a proper comparison of our previous
values for chondrites with those for
them, we calculated the values with the
ed counter efficiencies. Table 2 gives
ed tritium and argon 39 values for the
chondrites that we previously measured,

TABLE 2. Tritium and Argon 39 in the
Bruderheim and in Previously
Measured Chondrites

Tritium Decays/ kg min	Argon 39 Decays/kg min	H ³ /Ar ³⁹
260 ± 30	10 ± 1	26 ± 5
310 ± 30	8.5 ± 0.5	36 ± 6
210 ± 25	7 ± 1	30 ± 5
310 ± 60	5.3 ± 0.8	58 ± 15
	7.6 ± 0.2,	
	7.9 ± 0.2	
	8.7 ± 0.5,	
	9.2 ± 0.3	
	11.9 ± 0.5	
	7.1 ± 0.6	
	9.4 ± 0.5	

en and DeFelice [1960a].
en and DeFelice [1960b].
er, Schaeffer, and Davis [1960].

together with the Bruderheim results, and argon
39 values in chondrites obtained by *Stoennner*,
Schaeffer, and *Davis* [1960]. Table 2 shows that
there is not much variation in the activities of
the chondrites. The tritium values range from
200 to 300 decays/kg min; the argon 39 activi-
ties, from 5.3 to 12 decays/ kg min. The ratio
of tritium/argon 39 ranges from the 26 ± 5 for
Bruderheim to 58 ± 15 for St. Michel.

We crushed a sample of the Bruderheim me-
teorite, passed most of it through a 150-mesh
screen, and separated the highly magnetic from
the nonmagnetic particles. From visual inspec-
tion, we believe that the highly magnetic parti-
cles (10 per cent of the total weight) contained
about 20 per cent silicates. The argon 39 activity
of highly magnetic particles was 36 ± 6
decays/kg min, compared with the value $10 \pm$
1 in the whole rock samples and the value 4.8
 ± 0.6 in the nonmagnetic particles (see Table
1). The magnetic particles in Bruderheim have
the highest argon 39 content of any measured
meteoritic material. Table 3 lists the argon 39
values in metallic meteorites and in the metallic
phase of Bruderheim. The tritium content of
the magnetic particles, 90 ± 20 decays/kg min,
was low compared with the value 260 ± 30 for
the whole rock material and the value 290 ± 30
for the silicate particles.

TABLE 3. Argon 39 in the Metal Phase of Bruderheim and in Metallic Meteorites

Meteorite	Argon-39, decays/kg min
Bruderheim (metal phase)	36 ± 6
Sikhote-Alin*†	$4.3 \pm 0.2, 4.8 \pm 0.4$
Treysa†	13.3 ± 0.6
Pitts†	15.9 ± 0.5
Aroos‡	16 ± 2
Treysa§	21.6 ± 1.2

* *Fireman* [1958].

† *Fireman and DeFelice* [1960a].

‡ *Fireman and DeFelice* [1960b].

§ *Wänke and Vilcek* [1959].

Interpretation of results. Because this is the first time argon 39 has been measured in separated phases of a meteorite, we shall discuss this result first. Since the argon 39 content of the magnetic (metallic) phase of Bruderheim is significantly higher than that of metallic meteorites, and since the argon 39 and the tritium contents of the whole rock sample of Bruderheim were quite similar to those we previously measured in chondrites, we can conclude that chondrites in space are more affected by the cosmic-ray bombardment than metallic meteorites. Since the 325-year half-life of argon 39 is long compared with the period of the orbit of most bodies in the solar system, and very long compared with the duration of solar cosmic-ray events, the probable explanation for the high argon 39 content is that chondrites are less shielded while in space than metallic meteorites.

According to the chemical analysis [*Campbell and Baadsgaard*, 1961], the Bruderheim meteorite contains 22.5 per cent iron and 1.2 per cent nickel. The argon 39 content in the highly magnetic particles is consistent with the idea that the argon 39 (more than 70 per cent) is produced from the elements iron and nickel. The tritium was located principally in the silicate phase and was low in the metallic phase. In fact, the tritium content of the highly magnetic particles is so low that a large fraction of the tritium must have diffused out of the metal. The tritium content of the magnetic particles can be explained as resulting from a 20 per cent contamination with silicate material. This tritium loss is quite similar to that observed in most metallic meteorites.

Argon 37 has a 34 day half-life and can give information about cosmic-ray variations, both spatial (changes of the position of the body relative to the sun) and temporal (solar flares). The argon 37 activity in the whole rock sample, extrapolated to the date of fall, is twice the argon 39 activity. A similar result was obtained in the Hamlet meteorite. Since the ratio of argon 37 to argon 39 produced by energetic protons on iron is 0.8 ± 0.1 [*Schaeffer and Zähringer*, 1958] and since the ratio of argon 37 to argon 39 in Bruderheim is 2.2 ± 0.4 , there is an excess of argon 37 over what would be expected from the action of cosmic rays of constant intensity on iron. This excess argon 37 may have been produced from calcium (1.3 per cent abundant) by the cosmic-ray effects on Bruderheim may have been constant. Since the argon 37 was measured in the separated metallic or silicate phase, this question was not resolved.

It is fashionable at present to give exposure ages for meteorites. To obtain an exposure age for Bruderheim, it is necessary to combine a helium 3 measurement with our tritium value and an argon 38 measurement with our argon 39 result. Dr. Zähringer (private communication) has obtained the preliminary values of 50×10^{-6} cc/g of helium 3 and 1.4×10^{-8} cc/g of argon 38 for Bruderheim.

If we use these values in our calculations, we find that the ratio of helium 3 to tritium gives for Bruderheim an exposure age of 40 ± 10 million years, and the ratio of argon 38 to argon 39 gives 36 ± 4 million years. The measured values [*Goebel and Zähringer*, 1961; *Schaeffer and Zähringer*, 1958] 1.3 ± 0.3 for the production ratio of helium 3 to tritium, and 2.0 ± 0.2 for argon 38 to argon 39, are used in this calculation for the exposure age.

There are two interpretations of the short exposure ages of meteorites: space erosion, and recent large-body breakup. We favor the space erosion interpretation [*Whipple and Fireman*, 1959]. The exposure age of 36 m.y. corresponds to an erosion rate of approximately 10^{-6} cm/year for the Bruderheim meteorite.

Acknowledgments. We are indebted to Professor R. E. Folinsbee, who generously gave us a sample of Bruderheim. Professor C. Frondel kindly separated magnetic and nonmagnetic phases of the Bruderheim sample for us. We should like also

Professor F. L. Whipple for his constant encouragement.

REFERENCES

- n, F., J. Geiss, and D. C. Hess, Radiation from a meteorite from cosmic-ray-produced tritium, *Phys. Rev.* **107**, 540, 1957.
- n, F., P. Eberhardt, and D. C. Hess, ³⁹Ar Strahlungsalter eines Steinmeteoriten, *Naturforsch.*, **14a**, 500-503, 1959.
- l, F. A., and H. Baadsgaard, The Bruderheim meteorite chemistry and mineralogy, *Bull. Meeting AGU*, p. 56, 1961.
- l, E. L., Argon-39 in the Sikhote-Alin fall, *Nature*, **181**, 1613, 1958.
- l, E. L., and J. DeFelice, Argon 37, argon 39, and tritium in meteorites and the spatial constancy of cosmic rays, *J. Geophys. Research*, **65**, 3041-3044, 1960a.
- l, E. L., and J. DeFelice, Argon 39 and tritium in meteorites, *Geochim. et Cosmochim. Acta*, **24**, 183-192, 1960b.
- l, E. L., and D. Schwarzer, Measurement of ³⁹Ar, ³⁷Ar, and H³ in meteorites and its relation to cosmic radiation, *Geochim. et Cosmochim. Acta*, **21**, 252-262, 1957.
- Goebel, K., and P. Schmidlin, Tritium-Messungen an Steinmeteoriten, *Z. Naturforsch.*, **15a**, 79-82, 1960.
- Goebel, K., and J. Zähringer, Erzeugung von Tritium und Edelgasisotopen bei Bestrahlung von Fe and Cu mit Protonen von 25 GeV Energie, *Z. Naturforsch.*, **16a**, 3, 1961.
- Schaeffer, O. A., and J. Zähringer, Helium and Argon-Erzeugung in Eisentargets durch energiereiche Protonen, *Z. Naturforsch.*, **13a**, 4, 1958.
- Sprenkel, E. L., R. Davis, Jr., and E. O. Wiig, Cosmic-ray-produced Cl³⁶ and Ar³⁹ in iron meteorites, *Bull. Am. Phys. Soc.*, **4**, 223, 1959.
- Stoerner, R. W., O. A. Schaeffer, and R. Davis, Jr., Meteorites as space probes for testing the spatial constancy of cosmic rays, *J. Geophys. Research*, **65**, 3025-3034, 1960.
- Wänke, H., and E. Vilcsek, Argon 39 als Reaktionsprodukt der Höhenstrahlung in Eisenmeteoriten, *Z. Naturforsch.*, **14a**, 929-934, 1959.
- Whipple, F. L., and E. L. Fireman, Calculation of erosion in space from the cosmic-ray exposure ages of meteorites, *Nature*, **183**, 1315, 1959.

(Manuscript received July 7, 1961; revised July 28, 1961.)

On the Radioactivity of the Bruderheim Chondrite

M. W. ROWE AND M. A. VAN DILLA

*Los Alamos Scientific Laboratory, University of California
Los Alamos, New Mexico*

Abstract. A sample of the Bruderheim chondrite (fell March 4, 1960) has been examined with a γ -ray spectrometer, K^{40} , Al^{26} , Mn^{54} , and Na^{22} having been identified and quantitatively assayed. The concentrations found were 0.089 per cent potassium, 57 disintegrations per minute per kilogram Al^{26} , 82 disintegrations per minute per kilogram Mn^{54} , and 90 disintegrations per minute per kilogram Na^{22} , which are in good agreement with values found by other workers on this and other meteorites.

Introduction. The γ -ray spectrum of a 2.15-gram sample of the Bruderheim chondrite (fell March 4, 1960, in Alberta, Canada) has been investigated using a NaI (Tl) scintillation spectrometer technique previously described by La, Arnold, and Anderson [1960]; it is shown in Figure 1. The technique has been extended to make it more quantitative. The technique has been the proper way of accounting for scattering and absorption effects and for self-absorption. Our solution has been to make a thin, uniform layer of the same shape as the meteorite and to mix it with iron powder thoroughly mixed with an accurately known amount of the radioisotope to be quantitatively assayed.

Experimental procedure. The shell (Fig. 2) was prepared by covering the meteorite with a thin foil, pressed tightly against the surface to conform closely to its shape. A molding compound, Rezolin, Epoxy F (Rezolin L-933A, Epoxy Resin 'F', Rezolin, Inc., Santa Monica, California), is then painted on the aluminum-foil-covered meteorite in two halves with an unpainted layer about $\frac{1}{8}$ inch wide around the middle. After the compound hardens, the two halves are separated from the meteorite and then pressed together with more molding compound to form the completed shell. The composition of the molding compound, shown in Table 1, was chosen so that the resultant electron density was very close to chondritic; the average atomic number was high enough to make Compton effect the dominant interaction.

The Bruderheim composition used was that of Sogaard and Stelmach (private communication) (see Table 2). On this basis, the Bru-

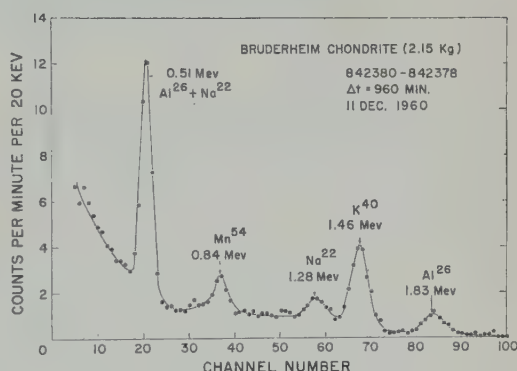


Fig. 1. γ -Ray spectrum of the Bruderheim chondrite.

derheim electron density was 2.92×10^{23} electrons per gram and that of the mock-up 2.85 to 2.88×10^{23} electrons per gram. After counting each mock-up in the same geometry as the meteorite, direct comparison of photopeak areas yielded the quantitative meteorite radioactivity.

Results. The Na^{22} and Mn^{54} solutions were both calibrated at the National Bureau of Standards with an accuracy estimated at ± 2 per cent. The Al^{26} was calibrated by comparing the positron annihilation peak of Al^{26} with that of the NBS-calibrated Na^{22} and also by comparing the 1.83-Mev peak of Al^{26} with a known 1.46-Mev K^{40} peak and making small corrections for photofraction, peak width, and self-absorption. The two results were within 7 per cent of each other. An accurately weighed amount of KCl constituted the potassium standard.

The Al^{26} content of the meteorite was also

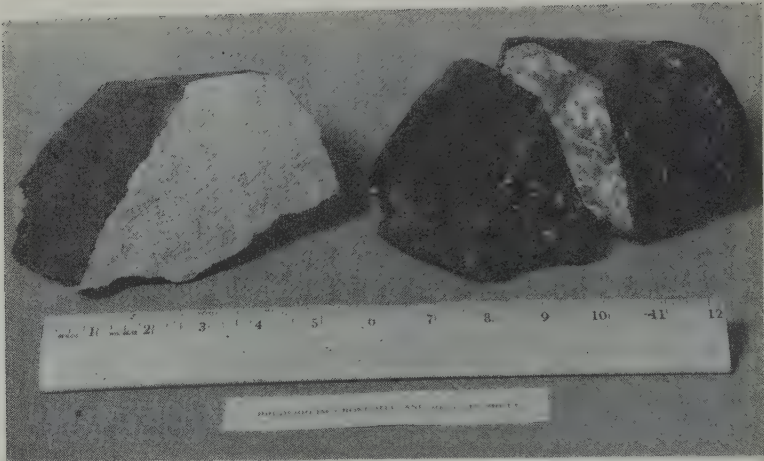


Fig. 2. The Bruderheim chondrite and Rezolin shell.

calculated by comparing its 1.83-Mev Al^{26} photopeak with its 1.46-Mev K^{40} peak. The three values for Al^{26} content by these methods, along with the results of the other measurements, are shown in Table 3.

Discussion. The spectrum of the Bruderheim chondrite is fairly complex, containing five γ -ray photopeaks at 0.51, 0.84, 1.28, 1.46, and 1.83 Mev. Of the four nuclides represented, Mn^{54} , Na^{23} , Al^{26} , and K^{40} , only the K^{40} is natural; the rest were produced by cosmic rays. Aluminum²⁶ is produced in chondrites by the spallation reaction $\text{Si}^{28}(p, 2pn)\text{Al}^{26}$ with contributions from $\text{Mg}^{26}(p, n)\text{Al}^{26}$ and $\text{Al}^{27}(p, pn)\text{Al}^{26}$. The Mn^{54} is probably produced mainly from the spallation reaction on iron, and the Na^{23} from the spallation on magnesium.

The decay of the 0.84- and 1.28-Mev peaks

has been followed for about 300 days to supply the identification. We found 200 days and 2 years compared with 291 days and 2.60 years [Strominger, Hollander, and Seaborg, 1958]. The fits were determined by an iterative least-squares computation using an IBM-704 curve-fitting code written by Moore and Zeigler [1960]. For further confirmation, the area of the 0.51-Mev peak was checked to be sure that Al^{26} and Na^{23} were the only significant contributors. This was found to be the case.

Since interpretation of the spectra becomes difficult with increasing numbers of peaks, it was considered desirable to process these data with an IBM-704 code written by Summers and Simpson (unpublished data, Kirtland Air Force Base, Albuquerque, New Mexico, 1960) as a check on the hand calculations. This code makes

TABLE 1. Composition of Four Mock-Ups for Bruderheim Meteorite Sample

	Potassium Mock-Up	Na^{22} Mock-Up	Mn^{54} Mock-Up	Al^{26} Mock-Up
Fe powder	1816 g	1726 g	1800 g	1537 g
KCl	162 g
NaCl	...	250 g	222 g	...
Dunite	500 g
Shell	172 g	174 g	128 g	113 g
Activity of standard	162 g KCl	0.0192 μC (NBS) on 1/16/61	0.0157 μC (NBS) on 11/18/60	3240 μC and 3470 μC (see text)
Electron density ($\times 10^{23}$ electrons per gram)	2.86	2.86	2.85	2.88

TABLE 2. Bruderheim Composition Used in Electron Density Calculations*

	Per Cent
SiO ₂	39.94
Al ₂ O ₃	1.86
Fe	8.59
FeO	12.94
FeS	6.38
MgO	24.95
CaO	1.74
Na ₂ O	1.01
NiO	1.30
Cr ₂ O ₃	0.60
MnO	0.33
TiO ₂	0.12
K ₂ O	0.13
P ₂ O ₅	0.29
H ₂ O ⁻	0.01
H ₂ O ⁺	0.10
CoO	0.05
C	0.04
Total	100.38

*sgaard and Stelmach, private communi-

a least-squares best fit between the experimental spectrum of the meteorite and a library of reference spectra (i.e., the mock-up spectra). The results of the two methods of calculation agreed very well.

The value we obtained for potassium content is in good agreement with one literature value but is in disagreement with the other. Baadsgaard is now checking his value (R. E. Folinsbee, private communication, 1961). Our value is also in good agreement with the average value of 0.086 per cent for chondrites [Edwards, 1955].

The Al²⁶ values are in good agreement with those reported in the literature. They also agree with average (51 disintegrations per minute per kilogram) for 14 chondrites (Rowe, Van Dilla, and Anderson, unpublished data).

The Mn⁵⁴ activity per kilogram total iron in Bruderheim is close to the value found for the Aroos siderite (fell November 24, 1959; Azerbaijan, USSR; 320-gram slice), as shown in Table 3. This is strong evidence that the Mn⁵⁴ in Bruderheim is produced by cosmic-ray reactions

TABLE 3. Radioactivity of the Bruderheim Chondrite

Hand Calculation	IBM-704 Code	Literature
0.089 ± 0.004 per cent	0.090 per cent	0.0896 ± 0.0023% 0.0046%*
57 ± 4 disintegrations/min/kg (internal comparison with K ⁴⁰)	58 disintegrations/min/kg	0.116 ± 0.008%† 60 ± 6 disintegrations/min/kg‡
57 ± 4 disintegrations/min/kg (Al ²⁶ standard calibrated with K ⁴⁰)		
61 ± 4 disintegrations/min/kg (Al ²⁶ standard calibrated with Na ²²)		
82 ± 7 disintegrations/min/kg	87 disintegrations/min/kg	100 ± 13 disintegrations/min/kg‡
362 ± 31 disintegrations/ min/kg Fe		470 ± 47 disintegrations/min/kg Aroos siderite§ 425 ± 40 disintegrations/min/kg Aroos siderite
90 ± 6 disintegrations/min/kg	92 disintegrations/min/kg	90 ± 10 disintegrations/min/kg‡ 89 ± 15 disintegrations/min/kg Brietscheid chondrite¶

*olds, J. H., private communication with R. E. Folinsbee.

sgaard and Stelmach, private communication.

la, Umemoto, and Arnold [1961].

Arnold, private communication.

erson, E. C., and M. A. Van Dilla, unpublished data.

ek and Wänke [1960].

on iron and that both meteorite samples were exposed to similar fluxes.

The Na^{22} concentration is in good agreement with literature values.

Acknowledgments. We are indebted to Dr. R. E. Folinsbee and Dr. D. A. Taylor of the University of Alberta, Edmonton, Alberta, Canada, for the 2.15-kilogram sample of this meteorite.

Work performed under the auspices of the U. S. Atomic Energy Commission.

REFERENCES

Edwards, G., Sodium and potassium in meteorites, *Geochim. et Cosmochim. Acta*, **8**, 285-294, 1955.

Honda, M., S. Umemoto, and J. R. Arnold, Radioactive species produced by cosmic rays in Järderheim and other stone meteorites, *J. Geophys. Research*, **66**(10), 1961.

Moore, R. H., and R. K. Zeigler, Los Alamos Lab. Rept. LA-2367, 1960.

Strominger, D., J. M. Hollander, and G. T. Seaborg, Table of isotopes, *Revs. Modern Phys.*, **30**, 585-904, 1958.

Van Dilla, M. A., J. R. Arnold, and E. C. Anderson, Spectrometric measurement of natural and cosmic-ray induced radioactivity in meteorites, *Geochim. et Cosmochim. Acta*, **20**, 115-121, 1956.

Vilcsek, E., and H. Wänke, *Z. Naturforsch.*, **15a**, 1004-1007, 1960.

(Manuscript received June 15, 1961.)

The Petrography and Chemical Composition of the Bruderheim Meteorite¹

MICHAEL DUKE, DONALD MAYNES, AND HARRISON BROWN

*Division of Geological Sciences, California Institute of Technology
Pasadena, California*

Abstract. The Bruderheim meteorite, a hypersthene chondrite, has been examined both petrographically and chemically. The minerals of the meteorite are olivine (Fo_{78}), hypersthene (Fs_{22}), plagioclase (Ab_{98}), apatite, merrillite, chromite, troilite, and metallic iron-nickel. Comparison of petrographic and chemical analyses shows that (1) pyroxene, troilite, and metallic iron-nickel are texturally related; (2) there is substantial CaO contained in the hypersthene; (3) there are important variations in samples the size of a thin section, and it is likely that there are also sampling errors on quantities of the size used for chemical analysis. The texture of the meteorite, including the principal chondrule types, is described.

Introduction. Through the kindness of Professor E. Follinsbee, this laboratory, together with many others, received a specimen of the Bruderheim meteorite for study. In view of the importance of this fall and in view of the fact that many workers are using it for a variety of purposes, it was deemed desirable to make an independent petrographic examination of the meteorite together with an independent chemical analysis. This work was purposely carried out entirely without knowledge of the results of previous workers.

One polished section and one polished thin section were prepared for microscopic examination. Minerals concentrated from a 1-gram sample by magnetic and heavy liquid techniques were of sufficient purity that refractive index measurements could be easily and accurately made for chemical analysis. Fifteen grams of the meteorite were chipped from a 60-gram fragment and crushed in a Plattner mortar to 80 mesh. These metallic fragments that flattened during the crushing process were separated by reduced mechanically to 80 mesh. A small portion of approximately 25 grams was chipped from the 60-gram fragment and reduced to 80 mesh. This portion was separated into 'magnetic' and 'nonmagnetic' fractions with a magnet. Separate splits of the powders were made for spectographic and X-ray fluores-

cence analysis. The results of the latter studies will be reported in a later paper.

Mineralogy. The nonopaque minerals of the meteorite are primarily olivine and pyroxene with minor amounts of feldspar, apatite, and merrillite. The opaque minerals are metallic iron-nickel, troilite and chromite.

Olivine is found as a major constituent of many chondrules and as grains of variable size in the groundmass. The refractive indices (NaD) of the olivine are: $n_x = 1.679 \pm .002$; $n_y = 1.696 \pm .002$; $n_z = 1.715 \pm .002$. The optic angle varies slightly between $2V = +88^\circ \pm 2^\circ$ to $-85^\circ \pm 2^\circ$, suggesting that there is a small variation in composition. The approximate composition of the olivine according to the index data of *Winchell and Winchell* [1951] is $\text{Fo}_{78}\text{Fa}_{22}$.

Orthopyroxene is abundant in chondrules and the groundmass. It is found as separate grains and as intergrowths with olivine. The pyroxene appears to partly replace olivine crystals in many cases. Some pyroxene grains contain numerous unoriented pyroxene inclusions that are not optically continuous with the host grain but are apparently of the same composition as the host.

The optic properties of the pyroxene are: $n_x = 1.678 \pm .002$; $n_y = 1.687 \pm .002$; $n_z = 1.693 \pm .002$; $2V = -85^\circ \pm 2^\circ$. These properties correspond to hypersthene with a composition $\text{En}_{78}\text{Fs}_{22}$ [Hess, 1960].

Plagioclase occurs as thin, irregular but coarsely crystalline intergrowths between olivine

grains in a few chondrules and also as interstitial material in the groundmass. Indices of the feldspar are: $n_x = 1.527$; $n_z = 1.538$; they do not vary more than ± 0.002 from these values. In thin section the plagioclase has low relief and low birefringence. It commonly forms a matrix surrounding fine granular pyroxene or olivine in cracks and at grain and chondrule boundaries. The optic angle of the feldspar varies widely from $2V = -50^\circ$ to $2V = +85^\circ$.

The optical properties of the feldspar suggest that it is a highly sodic plagioclase (Ab_{95}) of the high-temperature structural type. In experiments performed by *Smith* [1955] a large variation of optic angle was noted in samples showing incomplete transitions from the high- to the low-temperature structural type. This suggests that the range shown here is indicative of partial inversion from one type to the other.

Apatite ($Ca_5(PO_4)_3X$) and merrillite ($3CaO \cdot Na_2O \cdot P_2O_6$) were observed as anhedral crystals, generally free of inclusions, interstitial to the chondrules and the silicate groundmass. A small portion of these minerals is found in the chondrules. Apatite has $n_o = 1.656$, $n_e = 1.652$ and merrillite has $n_o = 1.624$, $n_e = 1.620$. They are of approximately equal abundance.

Troilite and metallic iron-nickel are present in small amounts in the chondrules, but the majority is present as irregular masses which are interstitial to the silicate portions of the groundmass. At the time of examination of the thin section and polished section, films of iron oxides were observed but it is likely that these were entirely due to atmospheric oxidation of the metallic phase.

Chromite is present in only small amounts. It is found predominantly in the silicate portions of the meteorite, including the chondrules and forms small, irregular grains which rarely reach a diameter of 1 mm.

Texture. The most conspicuous textural elements are the chondrules, which range from $\frac{1}{2}$ to 3 mm in diameter and have varying internal structures. Thirty-five chondrules are well developed in the thin section and perhaps as many again are present as poorly developed, indistinct, or fragmental chondrules. The well-developed chondrules were examined individually and grouped according to mineralogy. It was found that the mineralogical groupings of the chondrules correspond closely to their textural group-

ings; only three chondrules were seen that might be considered to be made up of two types, and the separate parts of these chondrules conform to the mineralogical groupings.

The three principal types of chondrules are mineralogically composed of (1) pyroxene, (2) pyroxene and olivine, and (3) olivine. The pyroxene chondrules are composed of 95 to 100 per cent hypersthene with only insignificant amounts of olivine. Scattered troilite and iron-nickel inclusions, about 0.02 mm in diameter, can be seen within the chondrules and somewhat larger and more abundant opaque grains occur near the borders. The chondrules are constructed of indistinct to well-defined sets of parallel pyroxene laths averaging 0.02 mm in width and of variable length. Complex intergrowths of lath sets produce a radiating appearance in some chondrules, although no pyroxene chondrule shows a well-developed point from which all laths radiate. With one exception the chondrules are not sharply defined and appear to grade into the groundmass. The exception is a well-developed pyroxene chondrule of distinct outline and made up of coarser laths.

The chondrules of the olivine-pyroxene type are composed of approximately equal amounts of olivine and hypersthene. These chondrules are characterized by a framework of olivine laths about 0.01 mm thick and up to 1 mm in length, a discontinuous rim of olivine crystals about 0.1 mm in diameter, and interstitial fine-grained pyroxene in the interior portions reaching 0.05 mm in grain size. The olivine in the core and rim is in optical continuity whereas the pyroxene is variable in optic orientation. There are a few opaque inclusions scattered throughout the chondrules.

The chondrules that are composed predominantly of olivine can be further classified on the basis of other phases that are present. The chondrules that contain small amounts of pyroxene and plagioclase are similar in structure to the olivine-pyroxene chondrules, but are different in that they have much coarser olivine laths and rims. A typical example has a parallel set of olivine laths approximately 0.1 mm in width that is in parallel orientation with a rim that is approximately 1 mm in width. Interstitial to the laths are anhedral pyroxene grains no more than 0.03 mm in size. The pyroxene

TABLE 1. Frequency of Chondrules in the Bruderheim Meteorite by Mineralogical Type

Mineralogical Type	Number
Olivine	13
Olivine and pyroxene	4
Olivine and plagioclase	14
Olivine and fine-grained aggregate	3
	1
Total	35

about 15 per cent of the chondrule, and a small amount of plagioclase is also present. In some of these olivine chondrules that contain small amounts of feldspar have very thick boundaries.

This type of olivine chondrule consists of an aggregate of randomly oriented olivine grains together with 5 to 40 per cent of fine-grained interstitial material composed chiefly of olivine. The fine-grained material is very similar in texture to the finest-grained portions of the groundmass. Opaque inclusions are abundant in these chondrules. The boundaries of the chondrules of this type are generally indistinct, and there are a few aggregates of olivine and fine-grained material present in the groundmass that may actually be chondrules, but cannot be identified as such.

Three chondrules were observed that cannot be placed uniquely into one of the groups. These chondrules are made up of olivine and pyroxene that exhibit the pyroxene-olivine texture and the coarse olivine structure. The olivine within these composite chondrules resembles the mineralogy in the manner of the major groups.

In addition, there is one chondrule that is made up by itself and is worth describing separately. It is quite round in cross section and is about 3 mm in diameter. The center is composed of an irregular intergrowth of pyroxene and olivine of about 0.5 mm average grain size. The outer one-half of the chondrule is composed of an aggregate of olivine of finer-grained texture than the center and which contains only small amounts of pyroxene and apatite. Close

to the rather sharp boundary of coarse- and fine-grained portions is a zone rich in iron-nickel grains with some troilite. The outer boundary surrounds a rim composed of pyroxene and troilite, containing only small amounts of metallic iron-nickel.

The thirty-five chondrules which were studied in the thin-section are tabulated by chondrule mineralogy in Table 1. The predominance of chondrules that fall into such a small number of groups suggests that there are at least three distinct mineralogical types of chondrules and that the textural and mineralogical variations are not due to random cutting of a single chondrule type by the thin section.

The groundmass of the meteorite is composed of a very inequigranular aggregate of olivine, pyroxene, feldspar, phosphates, and the opaque phases. The texture is in part fragmental, but there are many places where pyroxene can be seen to embay olivine grains. There is also a suggestion that the troilite and metallic iron-nickel are associated with pyroxene. Most of the large pyroxene grains in the groundmass are found near concentrations of opaque material. The feldspar is found in irregular patches where it is always observed to enclose numerous fine grains of pyroxene or olivine. The apatite and merrillite as well as most of the troilite and metal occur as irregular grains and granular aggregates that are interstitial to the silicate portions of the groundmass.

Modal analysis. A modal analysis was made

TABLE 2. Modal Analysis of the Bruderheim Meteorite

Mineral	Volume, %	Assumed Specific Gravity	Weight, %
Pyroxene	45.83	3.35	42.26
Olivine	42.83	3.45	40.67
Feldspar	0.50	2.6	0.36
Apatite and merrillite	0.43	3.0	0.35
Metallic	3.54	7.8	7.60
Troilite	5.08	4.8	6.71
Chromite	0.54	5.7	0.85
Holes	1.25	3.5*	1.21

* Holes weighted as average density of meteorite.

TABLE 3. Chemical Composition of the Bruderheim Meteorite

	1	2	3	Average	Magnetic	Non-magnetic	C. po
SiO ₂	39.56	39.54		39.55	15.51	43.31	39
TiO ₂	0.12	0.12		0.12	0.05	0.12	0
Al ₂ O ₃	2.15	2.15		2.15	0.98	2.36	2
Met. Fe	7.44*	7.18†		7.31	53.40‡	0.36‡	6
FeO + Fe ₂ O ₃ (as FeO)	13.80	13.55		13.89	6.90	14.97	13
MnO	0.32	0.32		0.32	n.d.	0.35	0
CaO	1.76	1.79		1.78	0.78	1.94	1
MgO	24.66	24.71		24.69	7.77	27.16	24
Na ₂ O	0.99	0.99		0.99	n.d.	1.07	0
K ₂ O	0.12	0.11		0.12	n.d.	0.12	0
H ₂ O ⁺	0.14	0.17		0.16	n.d.	0.18	0
H ₂ O ⁻	0.05	0.03		0.04	0.04	0.01	0
P ₂ O ₅	0.28	0.27		0.28	n.d.	0.30	0
FeS	6.56	6.59		6.58	2.21	7.25	6
Ni(as metal)	1.22	(1.39)§	1.22	1.22	9.36	0.12	1
Cr ₂ O ₃	0.53	n.d.		0.53	0.27	0.58	0
TOTAL	99.70	99.44		99.73	97.27	100.16	99
Total Fe	22.35	21.90§	22.23	22.29	60.16	16.62	22

* Riott's procedure.
† HgCl₂ procedure.
‡ Average of one determination each by Riott's and HgCl₂ procedures.
§ Discarded in averaging.
|| Cr₂O₃ from no. 1 is used.

by counting 2500 points on a 10·cm² area of the polished thin section in both reflected and transmitted light. The proportions of the non-opaque minerals were measured in transmitted light. The proportions of opaque minerals and the ratio of opaque to nonopaque minerals were measured in reflected light. The two measurements were combined to obtain the modal data which are shown in Table 2.

Chemical analysis. A number of the chemical procedures were adapted from those used by Lee C. Peck and J. J. Fahey at the U. S. Geological Survey.

Portions of the crushed sample that were used for the various determinations were separated with a sample-splitter, with the exception that grab-samples were used in the analysis of the 'magnetic' fraction.

A 1-gram portion was used for the determination of H₂O⁺, SiO₂, TiO₂, Al₂O₃, CaO, MgO, and total iron. H₂O⁺ was determined from the loss in weight at 105°C. After fusion with sodium carbonate, silica was recovered by a double dehydration with hydrochloric acid. The R₂O₃ group was precipitated with ammonia, the precipitate

was dissolved and aliquots were taken for various determinations. Silica present in ammonia precipitate was recovered and weight was added to that previously determined. TiO₂ was determined colorimetrically with hydrogen peroxide. Total iron was determined by reduction with silver and titration with potassium dichromate. Al₂O₃ was determined gravimetrically by precipitation with 8-hydroxyquinoline [see Flagg, 1948]. Calcium was precipitated twice as the oxalate, ignited and weighed as oxide. Magnesium was precipitated twice as phosphate, ignited, and weighed as pyrophosphate. The manganese present in the pyrophosphate was determined, and the weight was corrected for its presence.

Total water was determined by the Penfield method [see Hildebrand, Lundell, Bright, and Hoffman, 1953, pp. 827-828], with lead oxide as a retainer. The value for H₂O⁺ was obtained subtracting H₂O⁻ from total water.

Alkalies were determined in a 0.5-gram sample, essentially in the manner outlined by Brownock and Berthold [1953].

Manganese and phosphorous were determined

4. Comparison of Chemical Analysis of Bruderheim Meteorite with the Composition Calculated from the Modal Analysis

Element	Modal Analysis	Average Chemical Analysis
	40.39	39.55
	0.07	2.15
	12.22	13.89
	31.40	24.69
	0.18	1.78
	0.04	0.99
	0.17	0.28
	7.60	8.53
	6.71	6.58
	1.21	

of a solution of 0.5 gram of meteorite in a 1-gram portion after fusion with a carbonate-potassium nitrate flux. Manganese was determined colorimetrically as permanganate, sulphur gravimetrically by precipitation with barium chloride, and nickel gravimetrically by precipitation with dimethylglyoxime.

Nickel, chromium, and sulphur were determined in a 1-gram portion after fusion with a carbonate-potassium nitrate flux. Manganese was determined colorimetrically as permanganate, sulphur gravimetrically by precipitation with barium chloride, and nickel gravimetrically by precipitation with dimethylglyoxime.

Iron was determined by *Riott's* procedure and by the mercuric-chloride method described by *Lundell, Hoffman, and Riott* [1931]. Generally, *Riott's* procedure gives a higher result for metallic iron. Experiments with other meteorites have not yet indicated the cause of the difference.

Concentration of iron as oxides, reported by *Riott*, was obtained by subtracting the sum of iron as iron and iron as sulphide from total iron.

Analytical results are shown in Table 3. It is seen that the analysis is in satisfactory agreement with the composite result derived from independent analyses of the 'magnetic' and 'non-magnetic' splits.

Comparison of modal and chemical analyses. In making the composition of each of the phases to be that deduced by optical analysis and approximate chemical composition

can be deduced from the modal data. The results of such a calculation are compared with the chemical analysis in Table 4.

The calculated chemical analysis is similar to the chemical analysis but is divergent for a number of possible reasons:

1. The surface of the polished section is marked by a number of holes representing plucking out of grains during polishing. To the extent that some mineral may be preferentially plucked out, the modal analysis is in error.

2. The proportions of the minerals may not have been accurately determined microscopically. This in great part is due to the small size sample area, which is not precisely representative of the entire meteorite. It is also difficult to count accurately the tenuous feldspar grains and oxide films.

3. The presence of CaO and Al_2O_3 in the pyroxene and TiO_2 and Cr_2O_3 in either the pyroxene or chromite cannot be determined by optical methods. This will lead to an underestimation of the abundances of these oxides.

It should be noted that the proportions of troilite and metallic iron-nickel found in the modal analysis agree with those found chemically.

To attempt a further correlation between chemical and modal data a modified normative calculation was made on the silicate portion 'average,' 'magnetic,' and 'non-magnetic' chemical analyses. This calculation is different from the usual CIPW norm in that the FeO/MgO ratios of the pyroxene and olivine were chosen to be the values deduced from the optical properties of the minerals. The effect of this assumption is to free a certain amount of FeO for the oxide phases present in the meteorite and, in doing so, to make up for the lack of a figure for Fe_2O_3 in the chemical analysis. It is also assumed that all the Na_2O is combined as albite and that the excess CaO and Al_2O_3 are present in the pyroxene phase. This assumption is slightly in error because of the small amount of calcium in the plagioclase and sodium in the merrillite. The results are tabulated in Table 5.

A number of conclusions can be drawn from these calculations:

1. There is a concentration of pyroxene with the metallic phase in the magnetic fraction. This

TABLE 5. Modified Normative Calculation of Chemical Analyses

		Average	Magnetic	Nonmagnetic	
(KAlSi ₃ O ₈)	or	0.8	...	0.7	
(NaAlSi ₃ O ₈)	Ab	9.9	1.6	9.6	
(Mg, Fe, Ca, Al)(Al, Si)O ₃	py	41.9	86.4	42.8	
(Mg, Fe)SiO ₄	ol	41.2	...	40.7	Silicate phase re-
FeO·Cr ₂ O ₃	Cr	0.9	1.3	0.9	calculated to 100%
FeO·TiO ₂	Il	0.3	0.3	0.2	
3 CaO·P ₂ O ₅	Ap	0.7	...	0.7	
	FeO	4.2	10.5	3.9	
		100.0	100.1	99.9	
	MgSiO ₃	69.1	69.2	70.2	Pyroxene Compo-
	FeSiO ₃	21.2	21.8	21.7	sition recalculated
	AlAlO ₃	1.2	3.2	1.1	to 100%
	CaSiO ₃	8.5	5.9	7.0	
		100.0	100.1	100.0	

would be predicted from the distribution of pyroxene in thin section where it is observed that there tend to be concentrations of pyroxene around opaque grains in the groundmass. During the crushing process there would be a selective concentration of pyroxene in the metal phase due to embedding of the silicate grains.

2. There is a significantly greater amount of uncombined FeO in the magnetic portion which would correspond to concentration of oxidized metallic iron-nickel. Chromite is concentrated relative to silicates in the magnetic fraction.

3. The pyroxene probably does contain approximately 3 per cent CaO. This is unusual because orthopyroxenes do not have a high tolerance for CaO and because the feldspar would seem the more likely location of the CaO. The general similarity of the calculated CaO and Al₂O₃ for the pyroxenes of the three analyses strongly suggest that most of the CaO and some Al₂O₃ are in the pyroxene. The high Al₂O₃ content of the pyroxene calculated for the magnetic fraction is due to the lack of analyses for alkalis in this fraction. A value for Na₂O was chosen to eliminate a small amount of free SiO₂ in the modified norm. If there is approximately 0.4 per cent Na₂O in the magnetic fraction, the alumina content of the pyroxene would be the same as calculated for the other two analyses. Assigning more alkali, alumina, and SiO₂ to feldspar increases the proportion of olivine:pyroxene and thus increases the calculated CaO of the py-

roxene. The high value of calcium in the pyroxene suggests the possible presence of a di-sidic clinopyroxene. None was observed in the thin section.

4. By comparison of the norm with the modal analysis a fairly large sampling error is demonstrated on a sample of the size of the thin section. This is particularly true of the plagioclase which has a very irregular distribution, but it also appears to be the case with the free iron oxides. The effect of the underestimation of the components is considerable in the modal analysis.

5. When a weighted average is taken of the 'magnetic' and 'nonmagnetic' analyses, there appears to be a lower olivine:pyroxene ratio in the norm of the 'composite' analysis as compared to the 'average' analysis. It is therefore suggested that sampling errors can be noticeable on aliquots of the size taken for chemical analysis.

Acknowledgments. We are indebted to Professor R. E. Follinsbee for making a specimen of the Bruderheim meteorite available to us. We appreciate the critical advice and help given by Professor Leon T. Silver. Mr. R. von Huene prepared the polished section and the polished thin section. The help of Mr. Lee C. Peck and Mr. J. J. Fabian of the U. S. Geological Survey in establishing some of the analytical procedures used in this work is gratefully acknowledged. This work was undertaken with the help of grant no. NSG 56-60 given by the National Aeronautics and Space Administration and Atomic Energy Commission contract AT(11-1)-208.

REFERENCES

- , W. W., and S. M. Berthold, Contribution to geochemistry, 1949, *U. S. Geol. Survey* 92, part 1, 1-14, 1953.
- F., *Organic Reagents Used in Gravitational and Volumetric Analysis*, Interscience Publishers, New York, 164-165, 1948.
- H., Stillwater igneous complex, Montana. Quantitative mineralogical study, *Geol. Soc. Am.*, 80, 27, 1960.
- I. W. F., G. E. F. Lundell, H. A. Bright, I. Hoffman, *Applied Inorganic Analysis*, Wiley and Sons, New York, 1953.
- J. E., and M. G. Mellon, Colorimetric determination of phosphorus as molybdovanadoboric acid, *Ind. Eng. Chem., Anal. ed.*, 16, 1, 1944.
- Lundell, G. E. F., J. J. Hoffman, and H. A. Bright, *Chemical Analysis of Iron and Steel*, John Wiley and Sons, New York, 152, 1931.
- Riott, J. P. Determining metallic iron in iron oxides and slags, *Ind. Eng. Chem., Anal. ed.*, 13, 546-549, 1941.
- Smith, J. R., Effects of heating natural plagioclases, *Carnegie Inst. Wash. Year Book*, 55, 188-192, 1955.
- Williard, H. H., and L. H. Greathouse, The colorimetric determination of manganese by oxidation with periodate, *J. Am. Chem. Soc.*, 39, 2366, 1917.
- Winchell, A. N., and H. Winchell, *Elements of Optical Mineralogy*, part II, John Wiley and Sons, New York, 500 pp., 1951.

(Manuscript received August 4, 1961.)

Abstracts of the Papers Presented at a Symposium on Ground Water Portland, Oregon, November 16-17, 1960

Sponsored by the Section of Hydrology, American Geophysical Union

F. BERRY (Petroleum Research Corporation, Berkeley, Calif.), *Anomalous Low-Potential Applicability to Waste Disposal*. The most critical requirement for disposal of any liquid is that the waste must not be able to escape from the injection horizon. Escape would result in an unknown degree of contamination of other aquifers. The vertical hydrodynamic gradient is far more critical than the lateral gradient in preventing leakage of liquid waste. The potential that waste liquids be injected into a naturally existing low datum-pressure as well as with the datum-pressure in stratigraphic layers above and below the proposed disposal thereby insuring that all leakage would be to the designated injection horizon. Such low pressure (or low potential) zones may exist by means of 'short circuiting' of hydrodynamic flow controlled by gravitational forces and effects of membrane phenomena. Such short-circuiting phenomena resulting in either anomalous high or low potentials might be created by such paths as fracture permeability, fault system conformities, reefing, and exceptionally permeability sand channels. Such short circuiting exists in the Amarillo-Wichita Mountains in the Midale formation in southeastern New Mexico, in the Wheeler Ridge in the southern part of the Tertiary basins of the Snake River Valley, and in the Tertiary basins of Oregon and Washington. Other anomalous potential data occurring with anomalous salinities are explained by prior theories in hydrodynamic and geochemistry. Laboratory evidence indicates that compacted clay minerals act as semipermeable membranes and thereby exhibit pressure and salt-filtration effects. Three separated areas in North America (central Canada; San Juan Basin, New Mexico; Colorado; and Wheeler Ridge anticline, San Juan Valley) have anomalous potentials and concentrations that may be explained by the effect of water cross-formationally through acting as semipermeable membranes. Presumably salinity anomalies from other areas may also be explained by shale-membrane phenomena. Low-potential zones created either by the phenomena or 'short circuiting' systems make the one safe site for disposal of toxic waste within the earth's crust.

FISHEL (U. S. Geological Survey, Lawrence, Kans.), *Soil Water and Ground Water in Low Precipitation*. The amount of water retained in a soil column after it has drained to an

equilibrium condition will vary with height above the water table and with the texture of the material. The curvature of the air-water interface of the annular zones of water around the points of contact will result in a vapor pressure that will be in equilibrium with the surrounding vapor pressure and is a function of height above the water table. The curvature is proportional to its height above the water table; hence the amount of water retained in a soil column under equilibrium conditions will decrease with height. Water moves through the soil by gravity and capillarity, but at a low moisture content, in which the annular zones of water are no longer in direct contact with each other, the water moves by a process of evaporation from one interface and condensation on the adjacent interface of lower potential gradient.

WALTER H. GARDNER (Washington State University, Pullman, Wash.), *Water Movement in Soil above the Water Table*. Unsaturated flow of water in soil materials is governed by the equation, $v = -k\lambda \nabla \Phi$, where v is the volume rate of flow per unit cross section, k is the saturated permeability constant (included explicitly so that the equation reduces to the equation for saturated flow when $\lambda = 1$), λ is a channel factor which depends largely upon water content, and $\nabla \Phi$ is the gradient of the potential or moving force. Because of the difficulty of obtaining reliable potential measurements, a diffusion-type equation, $\partial\theta/\partial t = \partial/\partial x (D\partial\theta/\partial x)$, is often used where $\partial\theta/\partial t$ is the change in water content with time, $\partial\theta/\partial x$ is the change in water content with the position x , and D is a variable diffusion factor which must be obtained empirically. If the D in the diffusion equation is replaced by $k\lambda \partial\Phi/\partial\theta$, the diffusion equation and the potential equation are equivalent. Measurement of potential and problems associated with hysteresis of the water-tension curve are not involved explicitly when the diffusion approach is used. Practical implications of unsaturated flow theory are many. Most important is the fact that any type of porosity change encountered as water moves into a dryer soil results in slowing of the advance of the wetting-front. Clay layers, hard pans, and similar obstructions slow the advance because of resistance to flow in the very fine pores. Sands and gravels in finer materials temporarily stop the advance of a wetting-front because of the absence of small pores and channels which can be wet at the higher water tensions existing near the wetting-front. When the soil adjacent to such layers becomes very wet (tensions approaching zero)

then water can enter. The capacity of soil above such layers to retain water is often greatly increased over what would exist if the layers of coarse materials were not present. Research in unsaturated flow at the present time includes efforts to measure water potential more reliably, with temperature (particularly as involved at a wetting-front due to heat of wetting) being explicitly considered, efforts to develop practical infiltration equations, studies of the nature of the channel factor as used in the potential approach or the diffusion factor as used in the diffusion approach, and studies of practical applications of unsaturated flow theory to a large number of field problems—particularly where boundary conditions due to stratification or geometry of the water application are important. Work is under way at many universities and in federal research laboratories. At least one coordinated attack is being made: A project on 'Water Movement in Soil' has been initiated by the Western Regional Research Committee, W-68. Soil physicists in eight western universities are involved in this project.

PAUL H. JONES (U. S. Geological Survey, Idaho Falls, Idaho), *Storage and Movement of Water between Adjacent Surface- and Ground-Water Reservoirs*. The free interflow between adjacent surface-water and hydraulically interconnected ground-water reservoirs tends to reduce the amplitude of head fluctuations in either reservoir. This reduction is a consequence of the cyclic exchange of water between the surface and the ground. A part of the excess stream flow is generally salvaged as bank storage during flood stage. Certain geologic settings may favor the distribution of influent seepage from streams by allowing percolation through underlying deposits to distant ground-water reservoirs. The dry-weather flow of the streams may be in considerable part effluent seepage of the same bank-storage water. Differences in the quality of water in adjacent and interconnected surface- and ground-water reservoirs can result in costly treatment problems, and significant changes in the hydraulic characteristics of aquifers may occur as a consequence of water-temperature differences. Marked cyclic changes in head between the interconnected reservoirs may result in difficulties with structure foundations, or in water-control problems. Many techniques are available to the hydrologist for analysis of the interrelationships between interconnected surface- and ground-water reservoirs, but very few detailed investigations of the quantitative significance of interflow have been made. This aspect of hydrology is fundamental to any evaluation of river basin yield.

R. G. McMURTRY (U. S. Geological Survey, Missoula, Mont.), *Ground-Water Functions in the Yields of Type Stream Basins*. Forecasts of stream flow are becoming increasingly important and necessary for the efficient operation of the Pacific Northwest's hydroelectric power systems, as well

as for the most effective operation of irrigated and flood-control projects. The critical period, so far as power systems in the inland part of the Pacific Northwest are concerned, is from October through March, when stream flow may be maintained largely by ground-water discharge into stream channels. The discharge of ground water into surface channels (called inflow in this paper) is controlled mainly by the geology of the basin and the slope of the water table. There is a great difference in the total inflow per unit area in various basins and also a difference in inflow to each basin for different years; however, the variations in ground-water discharge per unit area in each basin (and from year to year in the same basin) do not reflect the differences in total inflow proportionally. These differences are not reflected proportionally because the capacity of the ground-water reservoirs to accept water is fairly constant in each of the basins, and the large variations in total inflow are transmitted from the basin through the surface-stream channels rather than through the ground-water reservoir. The percentage of total inflow represented by ground-water discharge in stream flow is a major factor in correlating ground-water levels with subsequent stream flow from the area. The time of availability of recharge has some effect on the ground-water discharge in stream flow. For example, the ground-water discharge per unit area in one of the tributary basins of the Columbia River during a dry year is greater than in a wet year. More irrigation water is applied in the dry years, so the ground-water reservoir is often filled to a slightly higher level in the years by seepage from irrigation water. The replenishment of a stream draining a basin in which there is extensive irrigation may differ from one in which there is little or no irrigation.

M. J. MUNDORFF (U. S. Geological Survey, Boise, Idaho), *The Water Yield of Drainage Basins, with Particular Reference to the Ground-Water Component*. The total perennial water supply available from any basin is the sum of the surface and underground outflow and the salvageable evapotranspiration. Thus, in arid basins it is important to evaluate each of these quantities as accurately as possible. How much of the perennial water supply can be intercepted and used depends on economic, legal, and political considerations in addition to hydrologic factors. It differs from place to place and varies with time. In deriving the water yield of a drainage basin, or in determining the ground-water component of yield, some type of water budget is made. In arid and semiarid regions a partial water budget is more accurate and gives as much usable information as a total water budget. Precipitation and surface flow can be measured directly. Ground-water outflow at some places can be measured directly, but it generally is obtained by indirect methods. Only indirect methods are used for determining evapotranspiration. A number of different methods and combinations

can be used to determine the total water and the amount of each component. One of determining total water yield is to compare similar hydrologic characteristics. For which the water yield can be measured or closely estimated, are used as index and the unknown water yields of similar are obtained from the correlation.

NELSON (General Electric Co., Richland, *Ground-Water Movement Rates*. A wide range of practical problems ranging from domestic supply through disposal of industrial effluents requires a knowledge of rates of groundwater movement. Obtaining the rate of groundwater movement seems easy, but it really represents the end product of a rather complex chain of analysis. Traditionally, such analysis has been limited to saturated flow in homogeneous media for which the Laplace equation applies. The hydrologist, however, is always faced with more complex systems than those described by the Laplace equation, and it is desirable to shift the emphasis from the special case to a broader class of equations capable of describing partially saturated flow in heterogeneous media. The more basic equation, utilizing the summation convention over double indices, is

$$\frac{\partial}{\partial X_i} \left(\frac{\partial \phi}{\partial X_i} \right) + \frac{\partial K}{\partial \phi} \left(\frac{\partial \phi}{\partial X_i} \right)^2 = \frac{\partial \theta}{\partial t} \quad (i = 1, 2, 3)$$

where ϕ is the capillary conductivity of the soil, θ is the piezometric head, i.e., sum of pressure and gravitational head.

The three space coordinates, as i takes on 1, 2, and 3.

where θ is the water fraction of saturation.

The more basic class of equations the Laplace equation is found as a special case. The greatest advantage in changing emphasis to the more basic class lies in the greater insight brought to even the special cases. Recent work at the University of Illinois has been concerned with devising numerical techniques and the associated computer programs to solve this class of nonlinear partial differential equations.

M. ROCKWOOD (U. S. Corps of Engineers, Portland, Ore.), *Infiltration from Snow*. The effect of ground conditions on snowmelt runoff were studied in the studies of the Cooperative Snowmelt Studies, in the monthly water balances for the laboratory areas, and in synthesis of the hydrographs. Soil moisture capacity in certain areas of western United States range from 4 to 6 inches of liquid water

which can be depleted by evapotranspiration and must be refilled (under snowmelt conditions) before runoff can occur. Ground-water storage, which effects a time-delay to runoff, was evaluated by standard recession analysis. Snowmelt rates during the spring period are generally low in comparison with storm rainfall, and over a 24-hour period average about .04 to .08 inch per hour, with maximum hourly rates of about .15 to .20 inch. In hydrograph reconstitutions, in the Salmon River Basin, Idaho (drainage area = 13,550 sq. mi.), for the spring snowmelt season, daily snowmelt over the snow-covered area is generally between 1.0 and 1.5 inches per day. Of this amount, 0.2 inch per day was considered to contribute to deep percolation, which appeared later as base flow. The remaining runoff excess from daily snowmelt was routed through surface and subsurface components of basin storage, in the proportion of 20 per cent surface and 80 per cent subsurface, as the best estimate of the snowmelt runoff entering the stream channels through surface runoff and that entering the zone of interflow.

M. I. RORABAUGH (U. S. Geological Survey, Tallahassee, Fla.), *Infiltration from Surface Water*. Although basic laws are well established for solution of the more simple problems associated with infiltration from surface water, practical solutions of the more complex problems having multiple interrelation of a large number of variables are not always achieved. Indirect methods using field data of limited accuracy sometimes compound errors to the point where the errors in final results are exceeded by the errors of component parts. Examples are: (a) determination of canal or stream seepage by subtraction of discharge measurements; (b) determination of lake or reservoir leakage by a water budget. Research is needed to develop theory and to translate the theory into usable techniques for problems of surface-water infiltration into multilayered aquifers, temperature variations, unsaturated flow beneath the surface-water body, and semipervious river or lake beds. Additional work is needed on better definition of evapotranspiration.

LEONARD SCHIFF (U. S. Department of Agriculture, Fresno, Calif.), *Ground-Water Recharge—Progress and Research*. Information has been published on soil and water treatments and operational procedures that increase both the hydraulic conductivity of soils and the hydraulic gradient. Hydraulic conductivity has been increased by the use of chemicals, organic residues, and grasses and by the disturbance or removal of soil layers of low conductivity. Hydraulic gradients have been increased by the use of greater depths of water on soil surfaces and in pits, shafts, and trenches. Soil clogging and clogging in model recharge wells have been alleviated or reduced by scraping and by the use of filters. Rates of infiltration into aquifer material have been increased by placing

various materials over the aquifer material. Systems including strip water spreading, to take advantage of lateral flow, and rotational water spreading are suggested. In the Ground-Water Recharge Research Project a recharge guide is being developed based on the work mentioned and experiments concerned with (1) establishing relationships between soil characteristics and water movements, (2) the feasibility of irrigation approaches to recharge, and (3) interactions between chemical characteristics of the applied water and soil and the resulting characteristics of the water table.

R. W. STALLMAN (U. S. Geological Survey, Denver, Colo.), *Significance of the Unsaturated Zone in Hydrology*. The interrelations between precipitation and surface runoff, between precipitation and ground-water recharge, and between release of underground storage and changing ground-water levels, are but a few examples in which the unsaturated zone exerts a controlling influence on water occurrence. Furthermore, the unsaturated zone dynamically affects water in the saturated zone. In contrast, in current hydrologic analyses the unsaturated zone is generally viewed as a static influence on water distribution. It is believed that the hydrologist's effectiveness in forecasting could be greatly improved through adopting a more realistic awareness of the unsaturated zone. Evidently this could be done today, for the laws defining flow through this zone, on a dynamic basis, are already known. The hydrologist need only apply them. However, research along the following general lines is needed to make such an approach feasible on a large scale: (1) Field methods for determining the hydraulic characteristics of the unsaturated zone over a wide range of fluid content. (2) Low-cost simplified computing procedures for forecasting flow conditions in the nonsteady state. (3) Field instrumentation for measurement of fluid content, liquid head, temperature, and salinity at depths up to several hundreds of feet below the land surface.

FRANK A. SWENSON (U. S. Geological Survey, Billings, Mont.), *Geohydraulics of Type Discharges to the Surface from Confined, Unconfined, and Perched Ground Water*. The Nation's subsurface storage of water, in the aggregate, far exceeds the storage available in surface reservoirs, and there is a growing awareness of the importance of subsurface storage in our complete water budget. Several very extensive, quite prolific artesian aquifers underlie much of central and eastern Montana. Though the Madison limestone is overlain by at least several thousand feet of other sediments in large areas, virtually all large springs and most large-capacity flowing wells obtain water from this formation. Discharge from the Madison takes place where upturned edges are cut by streams. In places water moves up along faults; the 190 ft³/sec discharged by Warm Springs and the 140 ft³/sec dis-

charged by Big Springs near Lewistown are good examples of this type. In other places water moves upward from the Madison along joints in overlying beds; the Giant Springs (flow 640 ft³/sec) near Great Falls are a prime example of this type of discharge. In places discharges of 10 to 15 ft³/sec from the Madison have been encountered in test wells. When extensive artesian aquifers discharge a considerable proportion of a stream's flow there will be little variation of flow in the stream. Even during prolonged droughts the tremendous reservoir backed up for many tens of miles does not become seriously depleted. Normally, it is difficult to increase the discharge of such aquifers appreciably. Unconfined aquifers are usually associated with unconsolidated alluvial deposits. In the Gallatin Valley of Montana unconsolidated deposits are locally more than 100 feet thick and a large quantity of water is stored in and discharged from them. Surface water applied for irrigation and downward percolation of water from streams and canals provide much of the recharge. At least 240,000 acre-feet of ground water enters and leaves this subsurface reservoir each year without any unnatural regulation. Additional water can readily be developed by means of drains and wells. In the aggregate, Montana has many square miles of irrigated stream terraces, but in a few places are they more than a mile in width. The ground-water reservoir provided by terrace deposits is therefore small when compared with valley fill and artesian aquifers. In general, perhaps ground-water bodies are not of broad extent, and the amount of water stored under these conditions is small. For full and complete utilization of our water resources we should plan on using not only our available surface reservoirs but also our vast subsurface reservoirs. For this we must have more detailed geohydrologic information.

C. V. THEIS (U. S. Geological Survey, Albuquerque, N. Mex.), *Dispersion Due to Geological Factors*. The dispersion of tracers or contaminants in a natural aquifer, both longitudinally and laterally, appears to be generally much greater than would be implied by theoretical or laboratory work on hydrodynamic or integranular dispersion. The permeabilities of the various beds in most detrital formations vary widely through a factor of 10 to 100 or even more. The implied comparable variation in velocity will result in a rapid longitudinal dispersion of any batch of tagged water placed in an aquifer and will greatly magnify the effects of integranular dispersion. The wide variation in permeabilities also inhibits vertical movement within the formation. Elongate lenses of high permeability must cause local anisotropy in the aquifer. Where lenses at different horizons are oriented differently, the water will not move directly down the overall hydraulic gradient, but will also rotate irregularly or otherwise move in a direction perpendicular to this gradient in cells the sizes of which will be dependent on the relative

ity, size, shape, and orientation of the masses. An originally compact mass of tracer drawn out over a wide cross section into thin streamers upon which intergranular and the dispersive effects of minor non-uniformities in the aquifer will be much more pronounced. By such processes widespread lateral dispersion can take place. Fluorescein dye placed in exceedingly permeable sediments at Hanford is dispersed longitudinally over a distance of 10 miles after 2 months' time and laterally over an angle of at least 30 degrees. Von Herzen's data on variation in tritium concentration in ground water at Carrizozo, N. Mex., also indicate rapid lateral and possibly longitudinal dispersion of batches of recharge carrying high tritium concentrations after bomb explosions. The indication of dispersion in nature has an important bearing on problems of radioactive and other waste disposal.

K. TODD (University of California, Berkeley, Calif.), *Limitations of Ground-Water Science*. Todd assumes that the science of ground water is incomplete, then research is needed to expand man's knowledge of this vital natural resource. Today's primary research needs in ground water involves factors listed below. They deserve attention and consideration of agencies and individuals interested in ground water. (1) Artificial recharge of ground water—maintenance of

high rates, geometries, and effects of nonhomogeneous formations. (2) Computers, analog and digital—their applications and utility in solving ground-water problems. (3) Dispersion, microscopic and macroscopic—magnitudes, controlling factors, and nonuniform flows. (4) Dissemination of ground-water knowledge—education in ground-water hydrology, awareness of problems, and benefits from related fields. (5) Economics of ground water—the cost of ground water, well costs, and the economics of overdraft. (6) Flow of ground water in relation to natural boundaries—effects of streams, topography, and recharge; velocity distributions; and actual movement. (7) Land subsidence by ground-water changes—types, causes, and the mechanics of aquifers. (8) Management of ground water—unified basin development, economic utilization, and conjunctive operation with surface water. (9) Nonhomogeneous aquifers—equivalent permeabilities, analytic treatment, and flow distributions. (10) Pollution of ground water—travel, detection, and control. (11) Radioactive isotopes in subsurface formations—movement, use as tracers, and effects on subsurface conditions. (12) Safe yield estimation—convenient methods for existing and proposed basin conditions. (13) Two-fluid flows underground—effects of density, mixing, and external influences; applications to connate and sea waters. (14) Subsurface exploration—development of convenient, economical, readily interpreted logging techniques.

Geomagnetic and Solar Data

J. VIRGINIA LINCOLN

Central Radio Propagation Laboratory
National Bureau of Standards
Boulder, Colorado

PRINCIPAL MAGNETIC STORMS

ance knowledge of the character of the records at some observatories as regards disturbances)

ry	Green- wich date	Storm-time		Sudden commencement			C- figure, degree of ac- tivity ⁴	Maximal activity on K-scale 0 to 9			Ranges			
		GMT of begin.	GMT of ending ¹	Type ²	Amplitudes ³			Gr. day	Gr. 3-hr. period	K- index	D	H	Z	
					D	H								Z
	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	1961	<i>h m</i>	<i>d h</i>		<i>'</i>	<i>γ</i>	<i>γ</i>					<i>'</i>	<i>γ</i>	<i>γ</i>
	Apr. 10	19 48	12 00	ms	11	5,6	6	178	1520	550
	Apr. 14	09 ..	15 23	ms	14	6	7	334	1710	1230
	May 5	16 21	7 00	ms	06	5	7	232	1660	1000
	May 7	02 ..	7 18	ms	07	4	6	94	940	610
	May 9	04 05	9 11	ms	09	4	7	130	1130	500
	May 25	02 11	25 21	s.c.*	+12	+55	+14	ms	25	4	7	210	1520	980
	Jun. 1	08 00	1 23	ms	01	4,5,6	6	120	1060	410
	Jun. 2	07 00	2 20	ms	02	4,5,6	6	208	1090	680
	Jun. 20	20 20	23 04	ms	21	4,6,7	6	207	1412	980
									22	4				
	Apr. 9	09 00	9 13	ms	9	4	7	105	450	530
	Apr. 13	14 51	15 21	s.c.	-11	+25	+4	ms	14	8	7	95	830	540
									15	1,2	7			
	May 6	05 00	6 15	ms	6	2,3,5	6	70	700	530
	May 25	02 00	25 16	ms	25	4,5	7	65	620	440
	Jun. 21	00 00	23 03	ms	21	4,5	7	80	960	800
									22	3	7			
	Jun. 29	02 00	29 12	ms	29	2	7	65	640	450
	Apr. 14	12 00	15 09	ms	14	7,8	7	55	335	210
	May 5	15 00	6 22	ms	5	6	6	20	180	60
	May 25	01 00	25 20	m	25	2,3,7	5	25	135	80
	Jun. 1	09 00	2 19	ms	1	6	6	20	185	70
	Jun. 20	16 17	23 04	s.c.	-1	+10	0	ms	21	5,8				
									22	1,7	6	40	255	115
	Jun. 29	00 00	29 22	ms	29	2	6	20	150	85
	Apr. 2	20 ..	3 12	ms	3	2	6	34	83	59
	Apr. 13	14 50	16 06	s.c.*	+4*	-34*	-3*	ms	14	8	6	51	135	226
									15	1,2				
	May 4	17 ..	9 13	m	4	8	5	18	120	104
									9	4				
	May 10	21 ..	13 19	m	13	1,2	5	21	90	38
	May 24	23 ..	26 04	m	25	2,3	5	20	100	84
	May 30	10	m	1	4,5	5			
	Jun.		03 12		2	4	5	19	104	48
	Jun. 20	18 ..	23 03	ms	22	1	7	37	157	146
	Jun. 29	00 10	29 12	s.c.*	0	+6*	-1	m	29	2	5	30	105	68

ate time of ending of storm construed as the time of cessation of reasonably marked disturbance movements in the
pecifically, when the K-index measure diminished to 2 or less for a reasonable period.
dden commencement; s.c.* = small initial impulse followed by main impulse (the amplitude in this case is that of
se only, neglecting the initial brief pulse); ... = gradual commencement.
mplitudes of D and Z taken algebraically; D reckoned positive if towards the east and Z reckoned positive if verti-
ls.
cribed by three degrees of activity: m for moderate (when K-index as great as 5); ms for moderately severe (when
for severe (when K = 8 or 9).

Observatory (Observer- in-Charge)	Green- wich date	Storm-time		Sudden commencement				C- figure, degree of ac- tivity ⁴	Maximal activity on K-scale 0 to 9			Range	
		GMT of begin.	GMT of ending ¹	Type ²	Amplitudes ³				Gr. day	Gr. 3-hr. period	K- index	D	H
					D (6)	H (7)	Z (8)						
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Tucson (C. J. Beers)	Apr.		1 15					m	1	1	5	12	83
	Apr. 2	20 ..	3 14					ms	3	2	6	13	112
	Apr. 9	06 ..	9 21					ms	9	4	6	9	44
	Apr. 13	14 50	16 03	s.c.	+2	+15	+1	ms	15	1,2,3	6	22	202
	May 1	23 ..	2 08					ms	2	1	6	14	99
	May 4	22 42	6 19	s.c.	+1	-12	-1	ms	6	2	6	22	82
	May 16	00 ..	17 02					m	16	2,3	5	12	55
	May 25	02 ..	26 03					ms	25	3	6	11	108
San Juan (M. Vazquez)	Jun. 20	18 ..	23 03					ms	22	1	6	16	178
	Jun. 29	29 ..	29 16					m	29	2,3	5	13	108
	Apr. 13	14 35	16 10	s.c.	+0.5	+13	-5	m	15	1	6	5	202
Honolulu (C. D. Upham)	May	None											
	Jun. 20	16 18	23 03					m	21	1,4,8	5	12	144
	Apr. 8	19 00	12 01					m	9	4	5	7	80
Huancayo (A. Giesecke) (N. Casaverde)	Apr. 13	14 50	15 24	s.c.*	-1*	+22	+9	m	14	8	5	13	185
	May 4	17 00	7 16					m	6	2	5	8	90
	May 22	01 37	23 07	s.c.	+2	+8	+4	m	23	2	4	10	50
	May 25	02 00	25 16					m	25	3	5	4	86
	Jun. 14	23 34	16 03	s.c.	+0	+20	+5	m	15	7	4	8	60
	Jun. 20	16 00	23 03					m	21	1	5	12	160
	Apr. 9	05 30	9 20					m	9	5	5	7	338
	Apr. 11	01 40	11 19					m	11	6,7	5	3	178
Port Moresby (J. A. Brooks)	Apr. 13	14 50	14 20	s.c.	+2	+67	+6	ms	14	8	6	8	314
	Apr. 14	23 26	15 20	s.c.	0	+41	+6	ms	15	1	6	8	313
	May 4	21 50	5 22					ms	5	6	6	6	255
	May 16	00 50	16 18					m	16	5,6	5	6	163
	May 25	01 52	25 19					m	25	6	5	6	123
	May 31	02 03	31 20					m	31	5,6	5	6	159
	Jun. 1	08 30	1 20					ms	1	6	6	7	202
	Jun. 7	05 35	7 21					m	7	5	5	5	124
	Jun. 20	16 18	22 22	s.c.	0	+24	+3	ms	21	5,8	6	8	288
	Apr. 13	14 50	15 07	s.c.	0	+31	+27	m	14	7,8	6	6	177
	May			No record 07-10 Hrs April 15									
	Jun. 20	16 18	23 00	s.c.	0	+6	+5	ms	21	4	6	5	158
Apia (J. G. Keys)	Apr. 9	06 ..	12 01					m	5		
	Apr. 13	14 50	16 03					m	5		
	May 4	16 ..	8 12					m	6	2	5	4	102
	May 25	01 56	25 15					m	25	3	5	2	108
	Jun. 20	16 ..	23 04					m	21	4,8	5	6	177
	Jun. 29	00 ..	29 16					ms	29	2	6	3	153
Gnangara (P. M. McGregor)	Apr. 14	00 ..	16 10					ms	14	8	6	28	178
	May 25	02 12	25 19	s.c.	+1	+6	+5	m	25	2,4,5,6,7	5	18	90
	Jun. 20	18 50	23 04					ms	21	8	6	22	126
Toolangi (C. A. van der Wall)	Apr. 14	00 00	15 24					m	14	6,7,8	5	28	236
	May	None							15	1	5		
Amberley (A. L. Cullington)	Jun. 20	20 ..	23 03					m	21	5,7,8	5	29	152
	Apr. 13	14 50	15 18	s.c.*	+0.6	+31	-10	m	21	1,4	5		
	May 6	Grad.	06 15					m	14	7	5	21	212
	May 25	Grad.	25 19					m	15	1,2	5	17	112
	Jun. 20	20 23	23 04	Grad.				m	6	2,3,4	5	15	126
	Jun. 29	00 00	29 12	Grad.				m	25	3,4	5	21	150
							m	22	3,4	5	17	106	
							m	29	2,3	5			

SELECTED GEOMAGNETIC AND SOLAR DATA

Kp, *Ci*, *Cp*, *Ap*, *K_F*, *Rz* and Selected Days

June 1961

Three-hour Range Indices Kp^2									Prel. ³ Ci	Cp^4	Ap^5	3-hr. Range Indices K_F ⁶		Prov. ⁷ Rz	
1	2	3	4	5	6	7	8	Sum				Values	Sum		
3o	3o	3o	5-	6-	5+	3+	3o	31o	1.3	1.3	30	3335	5434	30	30
4-	2+	4-	5-	5-	5-	5o	1o	30-	1.3	1.2	28	3345	4442	29	42
2+	4-	3+	3o	1+	1o	1-	2-	17o	0.6	0.6	10	3443	1212	20	48
2-	2+	3-	2+	3o	1+	1-	1+	15+	0.4	0.4	8	2333	3112	18	62
1+	3o	3-	2+	1+	2o	2+	2o	17o	0.4	0.5	9	2333	1232	19	55
2-	2-	3-	1+	2+	5+	4o	3o	22o	1.0	0.9	17	2232	3443	23	49
2o	3-	4-	4+	5o	4-	4o	3+	29-	1.1	1.1	24	2334	4343	26	40
4o	3+	3+	3o	3+	3-	3-	3-	25o	1.0	0.9	16	5442	3233	26	45
1-	1o	2o	2o	1+	2o	2+	2o	13+	0.3	0.3	6	1133	1232	16	58
1+	1+	2-	1+	2o	0+	0o	0o	8o	0.1	0.1	4	2222	2100	11	77
0+	1-	0+	0+	1-	1-	1o	0o	4o	0.0	0.0	2	0100	1021	5	82
1o	4-	3-	2o	2o	2o	1-	0+	14+	0.6	0.4	8	1432	2211	16	72
0+	1-	0+	1+	1o	1-	1-	1o	6o	0.1	0.1	3	1101	0122	8	75
1o	0o	1-	1o	0+	0+	1-	3+	7+	0.2	0.2	4	1012	0124	11	80
2+	2+	2o	3+	3o	3-	2+	2-	20-	0.7	0.6	11	2223	3232	19	123
4-	2o	1+	2+	2o	3-	3-	2-	18+	0.6	0.6	10	4222	2232	19	128
1+	2-	0+	1+	3-	1o	1+	2o	12-	0.3	0.3	6	2211	1122	12	128
1+	2-	2o	3o	3-	3+	3+	4+	22-	0.8	0.8	14	2232	2334	21	128
2+	3o	2o	2-	1o	1-	2o	2-	14+	0.3	0.4	7	3322	0122	15	112
2-	2+	2+	2-	0+	2o	3+	4o	18-	0.7	0.6	10	2322	0134	17	116
5o	5-	4+	6-	6-	4+	5o	7o	42-	1.7	1.7	58	5445	5346	36	128
7o	6-	5-	5o	4o	5o	5+	5o	42-	1.6	1.7	58	7555	4435	38	123
7-	2-	1o	1-	2-	3-	2o	2+	16-	0.4	0.5	9	3111	1223	14	96
2+	2o	1o	1o	1+	1+	2-	1+	12o	0.3	0.3	6	3310	1222	14	96
2-	2-	1+	3o	2o	2o	2o	3+	17o	0.5	0.5	9	2213	2123	16	72
1+	1o	0+	0+	3o	2o	2o	1-	11-	0.4	0.3	6	2100	3222	12	56
0+	2-	3o	3o	3-	1+	0+	0+	13-	0.4	0.4	7	1233	2111	14	51
0+	1o	1+	2-	2-	1+	1-	1o	9o	0.1	0.2	4	0221	1112	10	38
3+	6+	5-	3+	3-	2o	2o	2-	26o	1.4	1.2	25	3544	2222	24	59
2+	1+	1-	0+	0+	0+	1o	1+	8-	0.1	0.1	4	2210	0011	7	63
Means:									0.62	0.61	14			77.7	
No. of days:									30	30	30			30	

quiet days (Q), ten quiet days (Q or q), five disturbed days (D) selected by Committee on Characterization of Magnetic Disturbances, J. Veldkamp, Kon. Nederlandsch Meteorologisch Instituut, DeBilt,

magnetic planetary three-hour-range indices *Kp* prepared by Committee on Characterization of Magnetic Disturbances, J. Bartels, Chairman, University, Göttingen, Germany.

planetary magnetic character-figures, *Ci*, prepared by J. Veldkamp.

magnetic character-figures, *Cp*, prepared by J. Bartels.

geomagnetic amplitudes *Ap* (unit 2γ), prepared by J. Bartels.

Fredericksburg three-hour-range indices *K* ($K_9 = 500\gamma$); scale-values of variometers in γ/mm : $D = 2.7$; $Z = 3.0$; prepared by Robert E. Gebhardt, Observer-in-Charge, Fredericksburg Magnetic Observatory, Corbin, Virginia.

geomagnetic sunspot-numbers (dependent on observations at Zurich Observatory and its stations at Locarno and Arosa) prepared by M. Waldmeier, Swiss Federal Observatory, Zurich, Switzerland.

Letters to the Editor

The Bruderheim Meteorite

H. BAADSGAARD, F. A. CAMPBELL, AND R. E. FOLINSBEE

Department of Geology

G. L. CUMMING

Department of Physics

University of Alberta, Edmonton, Canada

Bruderheim, a detonating bolide, fell at 1:06 a.m., MST, on March 4, 1960. The fireball traveled on an azimuth of N 100°, at a slope of 40°, with an initial atmospheric velocity of 8 to 10 miles per second. More than 300 kilograms have been recovered from the fall area, a well-defined ellipse centered about latitude 53°54'N, longitude 112°53'W. This ellipse is 3½ miles long and 2¼ miles wide (5.6 by 3.6 km), with its long axis N80W and the larger individuals located near its southeast apex. Falling vertically onto frozen ground at terminal velocity, most individuals re-

bounded onto the snow, facilitating recovery. The largest individual weighs 31 kilograms, and the catalogued collection of individuals weighing more than 100 grams now totals 188.

Circumstances surrounding the fall and recovery have been detailed elsewhere [Folinsbee and Bayrock, 1961].

The stone selected for investigation was (Fig. 1), a large individual from the southeast apex of the fall. This stone originally weighed about 60 pounds, but was divided by the fire into a number of fragments. Twenty-five per-

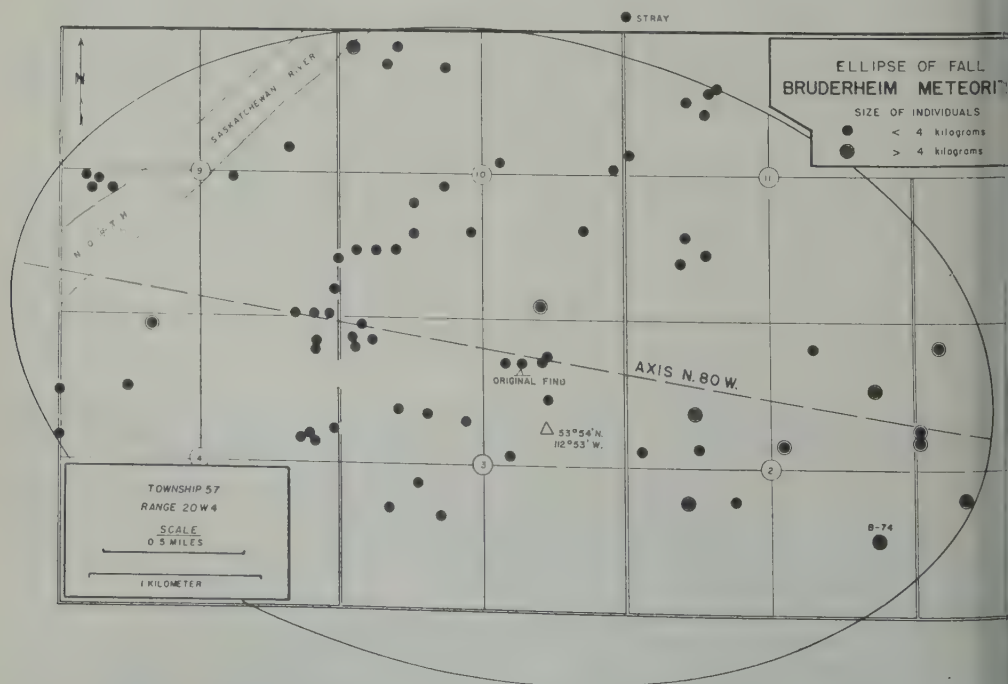


Fig. 1. Ellipse of fall, Bruderheim meteorite.

TABLE 1. Bruderheim Meteorite Samples Distributed for Investigation

of Investigator	Institution	Field of Investigation
J. R.	U. Calif., La Jolla	Cosmogenic isotopes
R.	Universität Bern	Rare gases
H. R.	Cal. Tech.	Chemistry and mineralogy
ue, J. D.	...	Nitrogen compounds
M.	U. Calif., Berkeley	Carbon compounds
A. J.	Mellon Institute	Gallium and germanium
n, W. D.	U. Kentucky	Neutron activation analysis
n, E. L.	Smithsonian Astrophys. Lab.	Ar ³⁷ , Ar ³⁹ , H ³
an, I.	U. S. Geol. Survey	Deuterium
P.	U. Minnesota	Strontium-rubidium dating
J.	Universität Bern	Rare gases
L. K.	CERN, Geneva	Ar ³⁹ , tritium
G.	U. Calif., La Jolla	Neutron activation analysis
vy, V. R.	U. Calif., La Jolla	Isotopic abundances and anomalies
B.	Fordham University	Carbon compounds
er, H. H.	Met. Investigations	Surface features
on, C. C.	Cal. Tech.	Lead isotopes
ds, J. H.	U. Calif., Berkeley	Rare gases, Ar ⁴⁰ /K ⁴⁰ dating
M. W. (and Van		
o, M. A.)	U. Calif., Los Alamos	Low-level γ -ray spectrometry
E. B.	U. Minnesota	Zinc
er, O. A.	Brookhaven	Rare gases
F. E.	U. S. Geol. Survey	Magnetic properties of crust
P.	U. Minnesota	Rare gases
H. E.	U. Calif., La Jolla	Rare gases
H. G.	McMaster University	Sulfur isotopes, rare gases
G. R.	Carnegie Inst. Wash.	Lead isotopes
ich, A.	U. Chicago	Radioactivity
ger, J.	Max-Planck Institut	Rare gases

of these were subsequently acquired by University, and, since the specimen was of value for museum purposes, and had been

recovered within a day of the fall, it has to date been used as the source of all Bruderheim investigation material (Table 1).

TABLE 2. Chemical Analysis and Normative Composition of the Bruderheim Meteorite

39.94	
0.12	<i>Normative composition</i>
1.86	Nickel-iron 9.94
8.59	Troilite 6.38
12.94	Olivine 41.65
6.38	Hypersthene 25.90
0.33	Diopside 5.34
24.95	Albite 8.52
1.74	Anorthite 0.17
1.01	Orthoclase 0.78
0.13	Chromite 0.92
0.29	Apatite 0.74
0.01	Ilmenite 0.21
0.10	
1.30	Analysis by H.
0.05	Baadsgaard and
0.60	A. Stelmach.
0.04	
100.38	

By means of a combination of gravity and magnetic methods a number of mineral separates have been obtained from the stone: olivine, hypersthene (bronzite), a pigeonitic pyroxene, plagioclase, kamacite, taenite, troilite, and chromite. These minerals are being analyzed. Some of the plagioclase is in the glassy form, maskelynite. Apatite or merrillite, carbon or hydrocarbons, ilmenite and lawrencite, though not separable, are probably present in small amounts.

Bruderheim appears to be a typical gray chondrite. Density of the meteorite, as determined with a quartz glass pycnometer, is 3.75, higher than the average for chondrites [Urey and Craig, 1953].

The chemical composition and normative mineral composition calculated from this analysis are given in Table 2. Chemically there is a remarkable similarity to Kyushu and Holbrook [Mason and Wükk, 1961]. In Prior's classification

TABLE 3. Bruderheim Meteorite Potassium-Argon Dates

No.	Sample	K ₂ O Per Cent	Ar ⁴⁰ /K ⁴⁰	Date
AK-139	Whole meteorite	0.134 (see Table 4)	0.20	1.
	25-100 mesh fraction	0.134	0.12	1.
AK-183	Feldspar	0.84	0.144	1.
	60-325 mesh			
AK-210	Calcium-rich pyroxene	0.22	0.15	1.
	170-270 mesh			
AK-252	Hypersthene	0.046	0.13	1.

* 2×10^{-8} cc Ar³⁶/g cosmogenic applied as correction.
† No correction for cosmogenic argon.
Constants used: $\lambda_{\alpha} = 0.589 \times 10^{-10}$ /yr, $\lambda_{\beta} = 4.76 \times 10^{-10}$ /yr.
 $K^{40}/K = 0.0118_1$ atomic per cent.

Bruderheim falls into the group of hypersthene-olivine chondrites of the Baroti type; in the classification of *Urey and Craig* [1953] it is closely comparable with their average analysis for the L group.

The K₂O content was measured by means of a flame photometer. The Na/K ratio is almost the same in the whole meteorite as in the feldspar (Table 3) and is closely comparable with the average Na/K ratio for chondrites of 7.9 obtained by *Edwards and Urey* [1955] and *Edwards* [1955]. However, our over-all alkali values are appreciably higher than those of Edwards, and for K in Bruderheim as established by γ -ray spectrography [*Rowe and Van Dilla*, 1961.]

No other elements show a significant variation from the averages.

Geiss and Hess [1958] have reported potassium-argon age determinations on a number of chondrites and other stony meteorites. The chondrites that they measure give ages between 4.0 and 4.4×10^9 years. *Reynolds and Iversen* [1957] report a potassium-argon age of 3.6 b.y. for Nuevo Laredo; *Stoener and Schriener* [1958] report potassium-argon ages to 10 b.y. for iron meteorites. A potassium-argon age determination of ≤ 1.9 b.y. was reported by *Thomson and Mayne* [1955] on the Bruderheim meteorite. This result on Monze is close to data for Bruderheim, both in potassium and diogenic argon content. *Signer* [1961] reports a 1.6-b.y. Ar⁴⁰/K⁴⁰ age for Bruderheim, supported by a He⁴/U age of 1.5 b.y. The writer has followed the suggestion of *Geiss and Hess* [1958]

TABLE 4. Alkalies on Bruderheim Mineral Separates

No.	Sample		K ₂ O	Na ₂ O	Na ₂ /K
R-31	Whole meteorite		0.141	1.01	7.9
R-31A			0.137	1.03	
R-32			0.133	1.00	
			0.125	1.01	
			0.137		
			0.128		
R-24	Whole meteorite <100 mesh	Average	0.134 ± 0.006 0.13	1.05	8.0
R-36	Purest feldspar	(flame)	0.97	7.72	8.0
		(gravimetric)	0.983		
R-37A	Olivine		0.05	0.46	9.2
R-37B	Calcium-rich pyroxene		0.27	2.05	7.6
R-45 (AK-252)	Hypersthene		0.046	0.31	6.0

argon-potassium determinations on the mineral separates from Bruderheim. The consistent Na/K ratios suggest that the various separates all the potassium is in the form of a fine-grained feldspar contaminant. The potassium content of the feldspar is relatively large, errors in potassium determination in this instance are not likely to be large. The 1.60-b.y. date is believed to be accurate, though subject to the uncertainties of feldspar dates by the $\text{Ar}^{40}/\text{K}^{40}$ method. These dates might be taken to support the conclusion of Mason [1960] that chondrites such as the Bruderheim form from carbonaceous chondrites close to the sun, and that Bruderheim is on this perilous journey about 1600 m.y. ago. [1961] criticizes Mason's paper, and offers alternative explanations of the low age of the Bruderheim, and Anders, 1960].

Bruderheim, a recently fallen meteorite which is available in quantity for study, is believed to be somewhat higher in potassium content than most chondrites, and to give a potassium-argon date.

Acknowledgments. We wish to acknowledge support from the National Research Council of Canada, the Geological Survey of Canada, and the University of Alberta. W. H. Johns, President of the University of Alberta, made a special fund available for the purchase of the Bruderheim meteorite.

REFERENCES

- Geiss, J., and H. C. Urey, Sodium and potassium in meteorites, *Geochim. et Cosmochim. Acta*, 8, 285-294, 1955.
- Geiss, J., and H. C. Urey, Determination of alkali metals in meteorites by a distillation process, *Geochim. et Cosmochim. Acta*, 7, 154-168, 1955.
- Folinsbee, R. E., and L. A. Bayrock, The Bruderheim meteorite—Fall and recovery, *J. Roy. Astron. Soc. Can.*, in press, 1961.
- Geiss, J., and D. C. Hess, Argon-potassium ages and the isotopic composition of argon from meteorites, *Astrophys. J.*, 127, 224-236, 1958.
- Goles, G. G., R. A. Fish, and E. Anders, The record in the meteorites, 1, *Geochim. et Cosmochim. Acta*, 19, 177-195, 1960.
- Mason, B., Origin of chondrules and chondritic meteorites, *Nature*, 186, 230-231, 1960.
- Mason, B., The origin of meteorites, *J. Geophys. Research*, 66, 2965-2970, 1960.
- Mason, B., and H. B. Wiik, The Kyushu, Japan, chondrite, *Geochim. et Cosmochim. Acta*, 21, 272-275, 1961.
- Reynolds, J. H., and J. I. Lipson, Rare gases from the Nuevo Laredo stone meteorite, *Geochim. et Cosmochim. Acta*, 12, 330-336, 1957.
- Rowe, M. W., and M. A. Van Dilla, On the radioactivity of the Bruderheim chondrite, *J. Geophys. Research*, 66 (10), 1961.
- Signer, P., Cosmogenic and radiogenic rare gases in chondrites (abstract), *J. Geophys. Research*, 66, 2560, 1961.
- Stoener, R. W., and J. Zähringer, Potassium-argon age of iron meteorites, *Geochim. et Cosmochim. Acta*, 15, 40-50, 1958.
- Thomson, S. J., and K. I. Mayne, The ages of three stony meteorites and a granite, *Geochim. et Cosmochim. Acta*, 7, 169-176, 1955.
- Urey, H. C., and H. Craig, The composition of the stone meteorites and the origin of meteorites, *Geochim. et Cosmochim. Acta*, 4, 36-82, 1953.
- Urey, H. C., Criticism of Dr. B. Mason's paper on 'The origin of meteorites,' *J. Geophys. Research*, 66, 1988-1991, 1961.

(Received July 17, 1961.)

Xenon in the Bruderheim Meteorite

W. B. CLARKE AND H. G. THODE

McMaster University, Hamilton, Canada

Gas samples used to determine the isotopic composition of xenon, obtained by vacuum melting samples of the Bruderheim meteorite, were analyzed in a 'static' mass spectrometer. The xenon content was accurately determined by means of calibrated quantities of Xe^{128} introduced into the high vacuum extraction system prior to melting. Samples of atmospheric xenon of approximately the same size as the meteoritic xenon samples were prepared from aliquots of air and used for comparison purposes.

The results of the mass spectrometric analyses are presented in Tables 1 and 2. Table 1 gives the content of 'atmospheric' xenon and 'excess' Xe^{129} , with the subtraction of atmospheric xenon from meteoritic xenon based on Xe^{130} . Table 2 gives a comparison between the isotopic abundances of meteoritic xenon and atmospheric xenon expressed in the following way:

$$\delta_i = \left(\frac{\text{Xe}^i}{\text{Xe}^{130}} \right)_{\text{meteorite}} - \left(\frac{\text{Xe}^i}{\text{Xe}^{130}} \right)_{\text{atmosphere}}$$

Table 2 shows that xenon from the Bruderheim meteorite exhibits the large excess of Xe^{129} first reported for the Richardton meteorite by Reynolds [1960a] and later by Signer [1960]. There is also the existence of definite secondary

TABLE 2. Comparison of the Isotopic Composition of Xenon from Bruderheim and Xenon from the Atmosphere

Isotope	δ_i
Xe^{124}	$+0.0074 \pm 0.001$
Xe^{126}	$+0.0101 \pm 0.001$
Xe^{128}^*	...
Xe^{129}	$+0.88 \pm 0.03$
Xe^{130}	$\equiv 0$
Xe^{131}	-0.19 ± 0.02
Xe^{132}	-0.37 ± 0.02
Xe^{134}	-0.11 ± 0.01
Xe^{136}	-0.0 ± 0.01

* A measurement of the isotopic abundance of Xe^{128} was impossible for sample III owing to memory from the two previous samples, I and II, spiked with Xe^{128} .

anomalies in the other isotopes similar to those observed for several other chondrites by Reynolds [1960b, c, d] as well as an extremely large anomaly in Xe^{130} . This fact, as well as the features, is in good agreement with recent work by Reynolds and Merrihue (private communication, 1961) on xenon from the Bruderheim meteorite.

TABLE 1. Xenon Content of the Bruderheim Meteorite

Sample No.	Description	Sample Weight, g	Xe of Atmospheric Composition, cc/STP/g	Excess Xe^{129} , cc/STP/g	$\text{Xe}^{129}/\text{Xe}^{130}$
I	20-80 mesh crushed 1 day before extraction	30.7	$9.9 \pm 0.3 \times 10^{-10}$	$4.5 \pm 0.4 \times 10^{-11}$	7.50 ± 0.1
II	Chunks, broken from the main piece 4 days before extraction	22.5	$10.0 \pm 0.3 \times 10^{-10}$	$4.56 \pm 0.15 \times 10^{-11}$	7.53 ± 0.1
III	Chunks, broken from the main piece 1 day before extraction	34.6	unspiked	unspiked	7.438 ± 0.001

REFERENCES

- J. H., Determination of the age of the
Phys. Rev. Letters, 4, 8, 1960a.
J. H., Isotopic composition of primordial
Phys. Rev. Letters, 4, 351, 1960b.
J. H., I-Xe dating of meteorites, *J. Geo-
search*, 65, 3843, 1960c.
- Reynolds, J. H., Isotopic composition of xenon
from enstatite chondrites, *Z. Naturforsch.*, 15a,
1112, 1960d.
Signer, P., Bestätigung des abnormal hohen ge-
haltes an Xe^{129} in Richardton, *Z. Naturforsch.*,
15a, 748, 1960.

(Received August 7, 1961.)

The Sulfur Isotope Abundances in
Abee and Bruderheim Meteorites

M. SHIMA AND H. G. THODE

McMaster University, Hamilton, Canada

Early work of Macnamara and Thode [1950] showed that whereas the S^{32}/S^{34} ratio for terrestrial samples varied by as much as 10 per cent, this ratio for meteoritic sulfur was remarkably constant. Since then, no significant variations in the S^{32}/S^{34} ratio for meteorites have been reported [Vinogradov, 1958, Ault and Kulp, 1959].

Recently, Thode, Monster, and Dunford [1961] reported the results obtained for seventeen meteorites. The results showed again the remarkable uniformity of the meteorites in sulfur isotope ratio. The maximum difference reported was of the order of $0.4\text{‰} \pm 0.1$.

The S^{32}/S^{34} ratios for the two Canadian meteorites, Abee and Bruderheim, have been compared with the average value obtained for seventeen meteorites [Thode, Monster, and Dunford, 1961]. In each case, separate determinations were made for the different phases present in the meteorite. Two methods were used in the preparation of the SO_2 samples for isotopic analysis. One involved the direct burn-

ing of sulphide in a stream of oxygen at temperature described by Thode, Monster, and Dunford [1961] and the other involved oxidation of sulphide with V_2O_5 , a method described by Gavelin, Parwell, and Ryhage [1960]. The results are reported in Table I.

REFERENCES

Ault, W. U., and J. L. Kulp, Isotopic geochemistry of sulphur, *Geochim. et Cosmochim. Acta*, **16**, 201-235, 1959.
Gavelin, S., A. Parwell, and R. Ryhage, Sulfur isotope fractionation in sulfide minerals, *Econ. Geol.*, **55**, 510-530, 1960.
Macnamara, J., and H. G. Thode, Comparison of the isotopic constitution of terrestrial and meteoritic sulfur, *Phys. Review*, **78**, 307-308, 1952.
Thode, H. G., J. Monster, and H. B. Dunford, Sulfur isotope geochemistry, *Geochim. et Cosmochim. Acta*, in press, 1961.
Vinogradov, A. P., Isotopic composition of sulfur in meteorites and in the earth, *Radioisotope Scientific Research*, **2**, 581-591, edited by J. Extermann, Pergamon Press, New York, 1960.

(Received August 7, 1961.)

TABLE 1. The S^{32}/S^{34} Ratios for the Abee and Bruderheim Meteorites

Sample	Method of Preparation of SO_2	S^{32}/S^{34} Ratio	$\delta\text{-}S^{34}\text{‰}$
1. Abee meteorite (chondrite, Alberta, 1952)			
Matrix part	Direct oxidation of Ag_2S (V_2O_5 oxidation)	22.227	-
	Direct oxidation	22.229	-
Fragment parts	Direct oxidation	22.232	-
	Direct oxidation	22.233	-
2. Bruderheim Meteorite (chondrite, Alberta)			
Metallic phase	Direct	22.230	-
	Direct	22.230	-
	V_2O_5	22.230	-
Mixed phase	Direct	22.228	-
Silicate phase	Direct	22.223	+
	V_2O_5	22.223	+
$\delta\text{-}S^{34}\text{‰} = \frac{(S^{34}/S^{32}) \text{ sample} - (S^{34}/S^{32}) \text{ standard}}{(S^{34}/S^{32}) \text{ standard}} \times 1000$			

The (S^{32}/S^{34}) standard in this case was the average for seventeen meteorites reported by Thode, Monster, and Dunford [1961].

Some Trace Element Abundances in the Bruderheim Meteorite

W. D. EHMANN, A. AMIRUDDIN, P. R. RUSHBROOK,¹ AND M. E. HURST*Department of Chemistry, University of Kentucky
Lexington, Kentucky*

elements tantalum, tungsten, and iridium were determined in the Bruderheim meteorite by means of neutron activation analysis. Tantalum was determined by counting the 115-day half-life Ta^{182} produced by (n, γ) reactions on 13 per cent naturally abundant Ta^{181} . Tungsten was determined by counting the 24-day half-life W^{187} produced by (n, γ) reactions on 26 per cent naturally abundant W^{186} [Amiruddin, 1961]. Iridium was determined by counting the 4.4-day half-life Ir^{192} produced by (n, γ) reactions on 38.5 per cent naturally abundant Ir^{191} [Rushbrook, 1960]. Gamma scintillation spectrometry was used to discriminate against background activities, and abundances were calculated by comparing sample activities with the activities of flux monitor foils which contained known amounts of the element to be determined. The foils were irradiated simultaneously with the meteorite, in the manner described by Ehmann and Huizenga [1959].

The abundances of tantalum, tungsten, and iridium in the Bruderheim meteorite, as determined in this work, are given in Table 1. They are consistent with data obtained for the same elements in other chondrites in this laboratory. The values listed for tantalum, however, must be regarded as only an order-of-magnitude determination, since only a preliminary run has been made on this sample and a very low chemical yield was obtained.

An interesting sidelight to the study of iridium abundances results from the fact that iridium is concentrated in the metallic phase of the meteorite. If a 100 per cent siderophilic character is assumed for iridium, and the metal-to-silica contents of a set of chondrite specimens are known, it is possible to calculate an average metallic phase iridium abundance from the gross iridium abundances. This was done

TABLE 1. Abundances of Tantalum, Tungsten, and Iridium in the Bruderheim Meteorite

	Abundance, ppm	Atomic Abundance (Si = 10 ⁶)
Tantalum	(0.04)*	(0.03)*
Tungsten	0.13	0.10
Iridium	0.51	0.40

* Preliminary result.

for a set of five chondrites whose metal-phase abundances were known. An average metal-phase iridium abundance of 2.85 ppm by weight was calculated. The individual values for the five specimens showed little deviation from this average. Using this value for the iridium content of the chondritic metal phase and the gross iridium content of the Bruderheim meteorite, a metal-phase abundance of 17.9 per cent was calculated for this meteorite. This calculation is certainly only approximate, since refinement in the value used for the average metal-phase abundance may result from further analyses. However, if the metal-phase iridium abundance is as constant as the few analyses done to date indicate, this technique could be applied to the determination of metal-phase abundances in small samples of stone meteorites, where magnetic separations would be difficult.

Acknowledgments. The support of the U. S. Atomic Energy Commission is gratefully acknowledged. The kindness of Dr. R. E. Folinsbee in supplying us with a specimen of the Bruderheim fall is greatly appreciated.

REFERENCES

- Amiruddin, A., Ph.D. dissertation, University of Kentucky, Lexington, 1961.
Ehmann, W. D., and J. R. Huizenga, Bismuth, thallium and mercury in stone meteorites by activation analysis, *Geochim. et Cosmochim. Acta*, 17, 125-135, 1959.
Rushbrook, P. R., M.S. thesis, University of Kentucky, Lexington, 1960.

(Received July 27, 1961.)

¹ Present address: General Electric Company, Washington.

Origin of Excess Xe^{129} in Stone MeteoritesP. M. JEFFERY¹ AND J. H. REYNOLDS²*Department of Physics, University of California
Berkeley, California*

Since the discovery of excess Xe^{129} in stone meteorites [Reynolds, 1960a], enough striking examples of the effect have been found [Reynolds, 1960b; Zähringer and Gentner, 1961] to establish beyond reasonable doubt that excess Xe^{129} is due to radioactive decay of extinct I^{129} . The usual interpretation has been that the I^{129} decay took place in situ, so that the 'formation interval' between the nuclear event that produced the I^{129} and the formation (or cooling) of a meteorite can be related to the $\text{Xe}^{129}/\text{I}^{127}$ ratio for the stone and to the 17 m.y. half-life of I^{129} . Some authors [Eberhardt and Geiss, 1960; Zähringer and Gentner, 1961], questioning this interpretation, have suggested that the I^{129} decay took place in a primordial planetary nebula where abnormal xenon may have been stored for long periods before meteorite formation. A crucial test of the first interpretation is whether the excess Xe^{129} is associated with the iodine-bearing minerals in a meteorite. This note describes an experiment that seems to prove this association for the enstatite chondrite Abee.

A 1.5-gram sample, consisting of three freshly broken interior chips, was irradiated in the Brookhaven pile for 480 hours at a flux of 6×10^{13} neutrons $\text{cm}^{-2} \text{sec}^{-1}$. The major effects of such an irradiation on the xenon mass spectrum are formation of Xe^{133} as an ultimate product of the (n, γ) reaction on I^{127} and formation of Xe^{131} in like manner from Te^{130} . There is no immediate production of Xe^{129} in the irradiation. After irradiation the sample was heated, in an outgassed system, at 1 hour for each temperature in 100° steps between 300° and 1500°C .

The 1500°C heating leaves the sample completely molten and outgassed. The gas released during each step was purified with hot Ti before the xenon fraction was admitted to an adjacent high sensitivity mass spectrometer [Reynolds, 1956] for isotopic analysis and xenon determination (by the peak height method). The results are plotted in Figure 1.

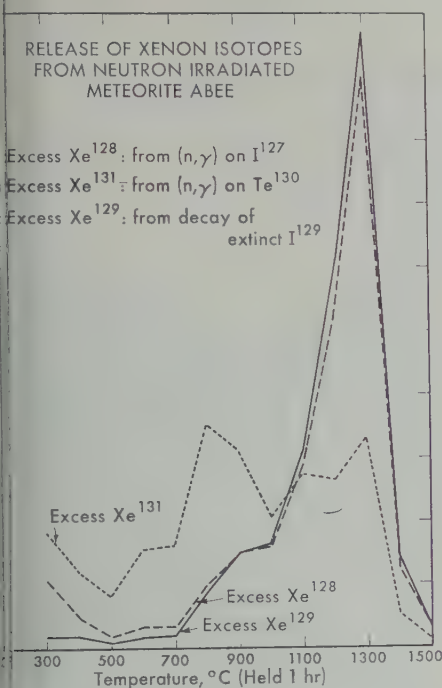
For simplicity of analysis the xenon was assumed to be a mixture of xenon of atmospheric composition, excess Xe^{128} from I^{127} , excess Xe^{131} from Te^{130} , and excess Xe^{129} from natural decay. The relative amounts of these components were, adding the yields at all temperatures, 1, 0.69, 0.10, and 1.08, respectively. The most striking feature of the results is the correspondence between the Xe^{129} and Xe^{131} patterns. They differ only in that there is a somewhat greater proportional release of Xe^{129} at the lowest temperatures. (It is not necessary to propose different sites for I and Xe^{129} to explain this effect, since any of three plausible mechanisms could be responsible: prior diffusion of Xe^{129} over the long lifetime of the meteorite, slight surface contamination of the specimen by terrestrial iodine, or recoil effect in the (n, γ) reaction.) The Xe^{129} pattern total yield is in good agreement with our previous work on an unirradiated specimen [Jeffery and Reynolds, 1961], indicating that the effects of radiation damage on these particular measurements are slight.

The contrast between the Xe^{131} release pattern and the others is particularly useful in interpreting the results. For example, escape of the gases from a melt can be invoked to explain only the data above 1300°C . We are forced to conclude from the release patterns below this temperature that I and Xe^{129} are at the same sites in the meteorite.

The tellurium is probably associated with the sulfide minerals in the meteorite. Golevsky

¹ Permanent address: Department of Physics, University of Western Australia, Nedlands, Western Australia.

² Presently in the Miller Institute for Basic Research in Science, University of California, Berkeley, California.



Release of xenon isotopes from neutron irradiated meteorite Abee.

[1961] have concluded that in two ordinary chondrites, Richardton and Bruderheim, tellurium and xenon reside in a fine-grained, soluble phase, tentatively identified as the 'tellurium phase.' In meteorite Abee the similar shape of the Xe¹³¹ release curve to the others below 800° suggests that the temperature release of Xe¹²⁸ and Xe¹²⁹ is also from the 'tellurium phase,' but the

bulk of the I and Xe¹²⁹ certainly resides elsewhere and in very retentive locations.

Because of the conclusions reached in the last two paragraphs, the I-Xe method of determining a formation interval appears, for Abee, to be on solid ground.

Acknowledgments. We wish to thank Mrs. Ruth-Mary Larimer and Mr. William Kwan for helpful assistance. We are grateful to Dr. Oliver Schaeffer and colleagues for cooperation in arranging for the Brookhaven irradiation. Dr. Heinz Stauffer independently suggested this experiment in a private communication which we wish to acknowledge. This research received supplementary support from the Institute of Geophysics, University of California and from the Alfred P. Sloan Foundation. The work was supported in part by the U. S. Atomic Energy Commission.

REFERENCES

- Eberhardt, P., and J. Geiss, Comment on the age of the elements, *Z. Naturforschg.*, **15a**, 547, 1960.
Goles, G. G., and E. Anders, On the chronology of the early solar system, *J. Geophys. Research*, **66**, in press, 1961.
Jeffery, P. M., and J. H. Reynolds, Concerning Xe¹²⁹ in the meteorite Abee, *Z. Naturforschg.*, **16a**, 431, 1961.
Reynolds, J. H., High sensitivity mass spectrometer for noble gas analysis, *Rev. Sci. Instr.*, **27**, 928, 1956.
Reynolds, J. H., Determination of the age of the elements, *Phys. Rev. Letters*, **4**, 8, 1960a.
Reynolds, J. H., Isotopic composition of xenon from enstatite chondrites, *Z. Naturforschg.*, **15a**, 1112, 1960b.
Zähringer, J., and W. Gentner, Zum Xe¹²⁹ in dem meteoriten Abee, *Z. Naturforschg.*, **16a**, 239, 1961.

(Received July 27, 1961.)

Cosmic-Ray-Produced V^{50} and K^{40} in the Iron Meteorite AroosHEINZ STAUFFER¹ AND MASATAKE HONDA

*School of Science and Engineering
University of California, San Diego
La Jolla, California*

The abundances of V^{50} and K^{40} in the iron meteorite Aroos have been determined by means of mass-spectrometric techniques. It was found earlier [Voshage and Hintenberger, 1959; Honda, 1959] that potassium separated from iron meteorites is enriched in K^{40} which is produced in cosmic-ray-induced spallation reactions. However, the absolute content of cosmogenic K^{40} has not yet been obtained accurately by mass-spectrometric measurements.

The content of vanadium in iron meteorites is unknown. Goldschmidt [1954] estimates a value of <5 ppm in the iron-nickel phase, while the troilite phase shows a considerably higher value. Recent studies on stable and radioactive cosmic-ray-produced nuclides [Arnold, Honda, and Lal, 1961; Honda and Arnold, 1961] permit an estimation of the abundance of cosmogenic V^{50} in the iron meteorite Aroos. Assuming a production rate of 100 atoms/min kg and an exposure age of 600 m. y. [Heymann and Schaeffer, 1961], we obtain a value of 3 ppb V^{50} . If the total vanadium content is <5 ppm, the enrichment of V^{50} should easily be detectable.

The experiments have been carried out on three different samples of Aroos. For the sample Aroos III, the procedure was as follows: About 15 grams of the meteorite were dissolved, using gaseous HCl and Cl_2 ; 2000 dpm of V^{49} tracer and 8.3 μ g of K^{39} (99.97 per cent) were added. The V^{49} tracer was prepared from a small iron target bombarded with 730-Mev protons. Its purity was checked by β - and X-ray counting and also by decay measurements over 10 months. The K^{39} was used as carrier and as spike for the isotopic dilution experiment.

After removal of iron by ether extraction, vanadium was extracted in $CHCl_3$, in the form of the cupferrate, and isolated by cation exchange. Potassium was separated by cation

exchange after electrolytic separation of niobium and other metals by means of a mercury cathode. The perchlorate solution of potassium was used directly for the mass-spectrometric measurements. The separated vanadium in perchloric solution was divided into three aliquots. The first aliquot was used for the determination of the chemical yield. The V^{49} activity was measured by means of a low-level X-ray proportional counter with a single-channel analyzer [Honda and Arnold, 1961]. The chemical yield was found to be about 40 per cent. The second aliquot was used for the determination of the isotopic composition of the extracted vanadium. A terrestrial vanadium spike was added to the third aliquot in order to determine the absolute amount of extracted vanadium. (In the experiment Aroos I an enriched V^{50} spike was added to a fourth aliquot, giving an independent check.)

The mass-spectrometric measurements were carried out with a single-focusing 60° deflection mass spectrometer with a 12-inch radius of curvature, built by Nuclide Analysis Associates. Ions were produced by surface ionization, using tantalum filaments. The ions V^+ and K^+ were measured. The sample size for a single run was of the order of 1 μ g for potassium and less than 0.05 μ g for vanadium.

The results of the potassium analyses are given in Table 1. The isotopic ratios are not corrected for the mass discrimination. However, the mass discrimination has been taken into account in the calculation of the absolute amounts of potassium, using Nier's values [Nier, 1950] as reference for the terrestrial potassium isotopic composition. The errors given include the statistical errors of the isotopic analyses and the uncertainty introduced by the correction for the mass discrimination. The absence of peaks at masses 42, 43, and 44 showed that there is no contribution of Ca^{40} and $Mg^{24}O^{16}$ to the K^{40} peak.

¹ Now at Physical Institute, University of Bern, Bern, Switzerland.

TABLE 1. Results of Potassium Measurements

Sample	K^{39}/K^{41}	K^{41}/K^{40}	K Contamination, ppm	Cosmogenic K^{40} , ppb
I	16.59 ± 0.05	365 ± 5	8.5 ± 0.5	0.54 ± 0.05
II	22.51 ± 0.05	100 ± 1	0.92 ± 0.03	0.49 ± 0.03
III, natural K	14.23 ± 0.10	565 ± 5		
IV, enriched K	2975 ± 30	24.0 ± 0.3		

Age cosmogenic $K^{40} = (0.51 \pm 0.03)$ ppb.

Errors of isotopic ratios = standard deviation of 15–20 spectra.

value for the K^{40} activity in Aroos and earlier [Honda and Arnold, 1961] is too compared with the present determination of K^{40} abundance. The earlier value, however, must be discarded, apparently because of a considerable loss of potassium during a long, free, chemical procedure. Honda and Schaeffer repeated the experiment using K^{42} tracer and K^{39} carrier. A value of 6.8 ± 0.7 dpm/kg was found for the cosmogenic K^{40} activity (published elsewhere). This result agrees with the limits of errors with the present spectrometric value, which corresponds to a value of 7.8 ± 0.5 dpm/kg.

The content of cosmogenic K^{40} can be compared with data on rare gases and radioactive elements in Aroos [Honda and Arnold, 1961; Schaeffer and Schaeffer, 1961]. The atomic ratio K^{40}/K^{39} is found to be 1.16 ± 0.12 . For the iron meteorites Grant and Williamstown values of 0.1 and 1.0 ± 0.2 were obtained earlier for the same ratio [Honda, 1959; Schaeffer and Schaeffer, 1960; Signer and Nier, 1960]. This suggests that the three meteorites had a similar formation history.

The results of the vanadium measurements are given in Table 2. Owing to the small sample sizes, ion beams of only about 5×10^{-16} A could be obtained. It was necessary to scan frequently over the masses 49 and 52 in order to correct for a possible contribution of Ti^{50} and Cr^{50} to the V^{50} peak. Corrections for Ti^{50} were found to be negligible, and the corrections for Cr^{50} were always smaller than 5 per cent except for the first run of Aroos I, where a 15 per cent correction had to be applied. White, Collins, and Rourke [1956] found a value of 398 ± 15 for the terrestrial ratio V^{51}/V^{50} , which is in good agreement with our value. Therefore no correction for mass discrimination has been applied. The errors given for the vanadium abundances include the mass-spectrometric errors in the determination of the chemical yield by V^{49} counting.

Since no blank experiment has been carried out so far, we do not know what fraction of the observed V^{51} abundance is primordial. The lowest measured value of 33 ppb is already more than a factor 100 smaller than the value given by Goldschmidt. The cosmogenic ratio V^{51}/V^{50}

TABLE 2. Results of Vanadium Measurements

Sample	V^{51}/V^{50}	Vanadium Abundances, ppb		
		V^{51}	V^{50}	Cosmogenic V^{50}
I	10.7 ± 0.7			
I + spike I	20.3 ± 0.9	55 ± 8	5.2 ± 0.8	5.1 ± 0.8
I + spike II	2.02 ± 0.03	52 ± 10	4.9 ± 0.9	4.8 ± 0.9
III	6.84 ± 0.15			
III + spike I	118 ± 3	33 ± 3	4.8 ± 0.4	4.8 ± 0.4
IV, terrestrial V	404 ± 2			
IV, enriched V^{50}	1.794 ± 0.010			

Age cosmogenic $V^{50} = (4.8 \pm 0.4)$ ppb.

Errors of isotopic ratios = standard deviation of 10–20 spectra.

TABLE 3. Relative Production Rates of Stable Cosmic-Ray Products in Aroos

Total Ne	A ³⁶	A ³⁸	K ⁴⁰	K ⁴¹	Sc ⁴⁶	V ⁵⁰
1.1	1.2	1.9	1.5	2.6	6	10

is expected to be about 3 (V⁵⁰ is a shielded isotope, whereas V⁵¹ measures the total production of the mass 51). Applying the correction for the cosmogenic V⁵¹ we get an upper limit of 20 ppb for the primordial vanadium abundance in Aroos.

The only stable, nonvolatile, cosmic-ray-produced isotope measured so far is Sc⁴⁶ [Wänke, 1958]. However, there is only one stable scandium isotope, and therefore the corrections for primordial and terrestrial scandium are uncertain. Since vanadium has two isotopes for comparison, an accurate correction for primordial and terrestrial vanadium can be applied. In fact, the results indicate that 99 per cent of the total measured V⁵⁰ is of cosmogenic origin.

Using a value of 600 m. y. for the exposure age of the Aroos meteorite [Heymann and Schaeffer, 1961], we obtain a value of 180 atoms/min kg for the V⁵⁰ production rate. Honda and Arnold [1961] measured the V⁴⁹ activity and obtained 164 ± 16 for the V⁴⁹ production rate. Thus the apparent production ratio V⁵⁰/V⁴⁹ is 1.1 ± 0.2 . Arnold, Honda, and Lal [1961] calculated 0.54 for this ratio.

In Table 3 the production rates of K⁴⁰ and V⁵⁰ are compared with those of other stable nuclides measured in Aroos [Heymann and Schaeffer, 1961; Signer, private communication]. The ratio K⁴¹/K⁴⁰ has been deduced from Voshage and Hintenberger's [1959] measurements in Treysa. The value for Sc⁴⁶ is obtained from the mean content measured in several meteorites [Wänke, 1960].

Among the stable nuclides studied so far in iron meteorites, V⁵⁰ is an example of a product of low-energy reactions, and neon is a typical high-energy product. Therefore the ratio V⁵⁰/Ne is a sensitive indicator for depth effects.

In stone meteorites the abundance of vanadium is too high (about 50 ppm) for detecting the enrichment caused by cosmogenic vanadium, but the separated metal phase might be suitable

for such measurements. All exposure ages determined in stone meteorites so far are based on rare-gas data. The possibility of diffusive losses cannot be excluded for samples with argon-potassium age much lower than 4.5 b.y. Accurate determinations of the cosmogenic V⁵⁰ and K⁴⁰ abundances could be used for estimating the extent of diffusive losses of cosmogenic rare gases. We are planning, therefore, to extend our measurements to the metal phase of stony meteorites.

Acknowledgments. We are indebted to Professor J. R. Arnold for suggesting this work and for critical discussions, and to E. L. Krinov for providing us with a specimen of the Aroos meteorite.

This work was supported by the National Aeronautics and Space Administration.

REFERENCES

- Arnold, J. R., M. Honda, and D. Lal, Records of cosmic-ray intensity in the meteorites, *J. Geophysical Research*, **66**, 1961.
- Goldschmidt, V. M., *Geochemistry*, Oxford, p. 4, 1954.
- Heymann, D., and O. A. Schaeffer, Exposure ages of some iron meteorites (abstract), *J. Geophysical Research*, **66**, 2535-2536, 1961.
- Honda, M., Cosmogenic potassium-40 in iron meteorites, *Geochim. et Cosmochim. Acta*, **23**, 148, 1959.
- Honda, M., and J. R. Arnold, Radioactive species produced by cosmic rays in the Aroos iron meteorites, *Geochim. et Cosmochim. Acta*, **25**, 219, 1961.
- Nier, A. O., A redetermination of the relative abundances of the isotopes of carbon, nitrogen, oxygen, argon, and potassium, *Phys. Rev.*, **78**, 789, 1950.
- Schaeffer, O. A., and J. Zähringer, He, Ne, and Ar isotopes in some iron meteorites, *Geochim. et Cosmochim. Acta*, **24**, 94, 1960.
- Signer, P., and A. O. Nier, The distribution of cosmic-ray-produced rare gases in iron meteorites, *J. Geophys. Research*, **65**, 2947, 1960.
- Voshage, H., and H. Hintenberger, Kalium als Reaktionsprodukt der kosmischen Strahlung in Eisenmeteoriten, *Z. Naturforsch.*, **14a**, 828, 1959.
- Wänke, H., Scandium-45 als Reaktionsprodukt der Höhenstrahlung in Eisenmeteoriten, I and II, *Z. Naturforsch.*, **13a**, 645, 1958; **15a**, 953, 1960.
- White, F. A., T. L. Collins, and F. M. Rourke, Search for possible naturally occurring isotopes of low abundance, *Phys. Rev.*, **101**, 1786, 1956.

(Received July 13, 1961.)

The Stability of the Interface between the Solar Wind and the Geomagnetic Field

A. J. DESSLER

*Lockheed Missiles and Space Company
Palo Alto, California*

The purpose of this letter is to present evidence that suggests that the interface between the solar wind and the geomagnetic field is stable rather than turbulent. The argument that the interface shows no gross instability is based on examination of transient geomagnetic activity and magnetometer data obtained at the earth's surface. Much of the observed geomagnetic activity is due to local ionospheric current systems (tidal and auroral currents). It has been suggested that some of the transient surface fluctuations in the geomagnetic field are transmitted across the interface between the solar wind and the geomagnetic field to the earth's surface by hydromagnetic waves [Dessler, 1958]. It is proposed here that such hydromagnetic waves are excited by a varying energy density in the solar wind instead of by turbulences or a gross instability at the interface.

It is generally observed on magnetometer records that the period immediately following the sudden commencement of a geomagnetic storm (the initial phase) is not disturbed to any appreciable degree greater than the period immediately preceding the sudden commencement. There is, however, a temporary 'ringing' of the geomagnetic field that is often observed at the onset of the SC [Campbell, 1959; Benioff, 1960]. Two examples of a quiet period following the initial phase are shown in Figure 1.

The initial phase of a geomagnetic storm is attributed to the pressure of solar plasma on the geomagnetic field. Therefore, the initial phase is assumed, with some certainty, to be a function of time when the solar wind is flowing across the geomagnetic field with a velocity $\sim 10^8$ cm/sec [e.g., see Chapman and Ferraro, 1940; Francis, and Parker, 1960]. If the interface between this solar plasma and the geomagnetic field were grossly unstable, large-amplitude hydromagnetic waves would be generated

by the turbulent motion of the geomagnetic field lines, and waves would propagate to the earth's surface where they would be observed as fluctuations in the geomagnetic field. Figure 1 shows two cases where this effect does not occur. Of course, cases can be found where the geomagnetic field at the earth's surface becomes highly disturbed immediately following the SC and all through the initial and main phases that follow. However, the fact that such cases as shown in Figure 1 may be easily found among magnetometer records argues against a natural gross instability at the boundary between the geomagnetic field and the solar wind.

The comparison of magnetic disturbance amplitude before and after an SC may be extended to much shorter period fluctuations than would be possible with standard magnetometers by utilizing data from instruments especially designed to record short-period magnetic fluctuations. Examination of results covering periods of 5 to 30 seconds [Campbell, 1959], 0.3 to 120 seconds [Benioff, 1960, and private communication], and 0.1 to 50 seconds (L. Tepley, private communication) confirm the conclusion reached with standard magnetograms such as shown in Figure 1. Attenuation of hydromagnetic waves with periods shorter than about 1 second may be significant [Francis and Karplus, 1960]. However, the apparatus used for the short-period measurements had threshold sensitivities of less than 0.01 γ , so that, multiplying this sensitivity by the maximum calculated attenuation (factor of 20 at 1 cps), we find that the wave amplitude for periods near 1 second must be less than 1 γ above the ionosphere. If the field were grossly turbulent, we might expect waves to be generated with an amplitude ΔB of the order of the value of B at the interface. If we assume that the interface has pushed in to 8 earth radii immediately following an SC, a value $\Delta B \approx 100 \gamma$

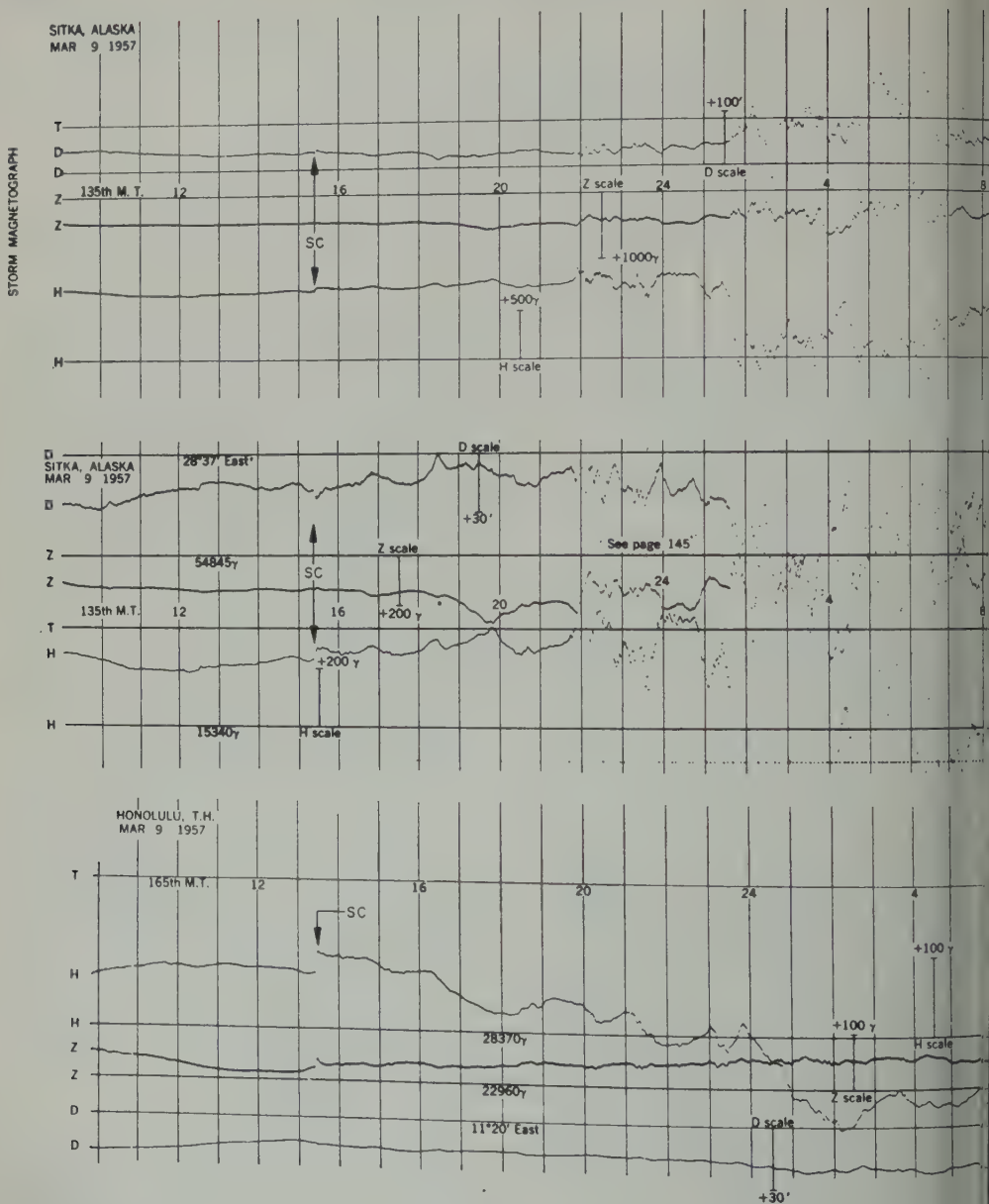


Fig. 1a

is obtained. The combination of geometric and dissipative attenuation must then be unreasonably large (i.e., greater than 10^3) in order to account for the absence of magnetic activity often observed at the earth's surface during the initial phase (e.g., Fig. 1). And, if attenuation arguments were used to explain the comparative absence of magnetic activity during the

period directly following the SC, how then the sudden commencement phenomenon and intense world-wide activity that occurs during the main phase to be explained? Thus we are led to the conclusion that the flow of plasma past the geomagnetic field does not lead to turbulence and the subsequent generation of large-amplitude hydromagnetic waves.

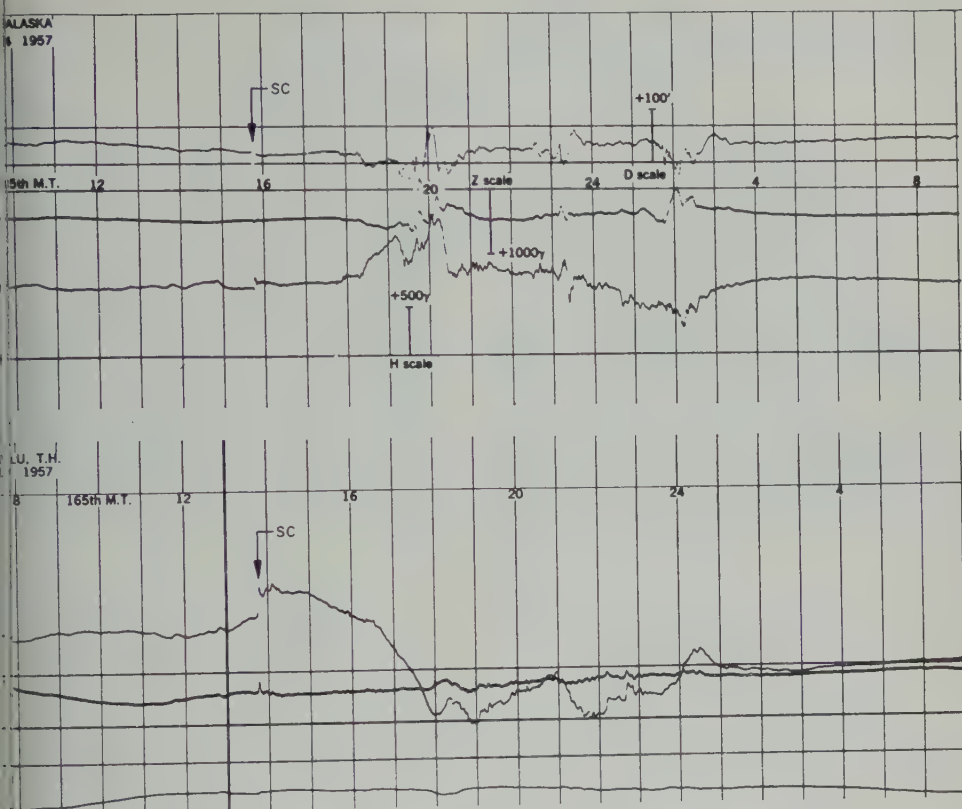


Fig. 1b.

Fig. 1, *a* and *b*. Magnetograms from Honolulu, Hawaii, and Sitka, Alaska, for two magnetic storms. The sudden commencements are marked SC. Note the magnetically quiet periods following these SC's; it is these quiet periods which are presented as evidence that the boundary between the geomagnetic field and the solar wind is stable.

ade ΔB even approaches the value of B at the interface.

On the basis of theoretical calculations, Dunlap [1955] and Parker [1958] have concluded that the interface is unstable. Also, satellite and probe magnetometer measurements have been interpreted as indicating instability and turbulence at the limit of the geomagnetic field [e.g., Sonett, Judge, and Smith, 1960; Smith, and Sims, 1960]. Conversely, theoretical investigations by T. Northrop (private communications), D. Beard (private communications), and M. Vallarta (private communications) suggest that the surface is stable. It is not the intent of this letter to reinterpret past work as has been indicated that the boundary should be stable; rather, the purpose is to point out that surface magnetometer data appear to show

that the interface between the solar wind and the geomagnetic field is stable in a gross sense (small-scale turbulence such that $\Delta B \ll B$ at the interface cannot be ruled out by the evidence presented here). Those world-wide fluctuations in the geomagnetic field that are transmitted from the interface to the earth's surface by hydromagnetic waves must then be due solely to energy-density fluctuations in the solar wind. If this conclusion be true, theories of aurora, Van Allen radiation, or magnetic storms that utilize the concept of turbulent solar injection in an important way, must be re-examined.

Acknowledgments. This work was supported by the Geophysics Research Directorate of the Air Force Cambridge Research Laboratories under contract AF 19(604)-7989.

I would like to thank Dr. Hugo Benioff for mak-

ing his ELF magnetic data available for this work, and Dr. L. Tepley for information regarding his data analysis on the sudden commencement ELF spectrum.

REFERENCES

- Benioff, H., Observations of geomagnetic fluctuations in the period range 0.3 to 120 seconds, *J. Geophys. Research*, **65**, 1413-1422, 1960.
- Campbell, W. H., Studies of magnetic field micro-pulsations with periods of 5 to 30 seconds, *J. Geophys. Research*, **64**, 1819-1826, 1959.
- Chapman, S., and V. C. A. Ferraro, The theory of the first phase of a geomagnetic storm, *Terrest. Mag. and Atmospheric Elect.*, **45**, 245-268, 1940.
- Coleman, P. J., Jr., C. P. Sonett, D. L. Judge, and E. J. Smith, Some preliminary results of the Pioneer V magnetometer experiment, *J. Geophys. Research*, **65**, 1856-1857, 1960.
- Dessler, A. J., Large amplitude hydromagnetic waves above the ionosphere, *J. Geophys. Research*, **63**, 507-511, 1958; and *Phys. Rev. Letters*, **1**, 68-69, 1958.
- Dessler, A. J., W. E. Francis, and E. N. Parker, Geomagnetic storm sudden-commencement recurrences, *J. Geophys. Research*, **65**, 2715-2719, 1960.
- Dungey, J. W., Electrodynamics of the outer atmosphere, *The Physics of the Ionosphere*, T. G. Fox, ed., Physical Society, London, pp. 229-236, 1955.
- Francis, W. E., and Robert Karplus, Hydromagnetic waves in the ionosphere, *J. Geophys. Research*, **65**, 3593-3600, 1960.
- Parker, E. N., Interaction of the solar wind with the geomagnetic field, *Phys. Fluids*, **1**, 171-176, 1958.
- Sonett, C. P., E. J. Smith, and A. R. Sims, Survey of the distant geomagnetic field: Pioneer I and Explorer VI, in *Space Research*, edited by J. H. Kallmann-Bijl, North Holland Publishing Company, Amsterdam, pp. 921-937, 1960.

(Received August 3, 1961.)

Note on Hydromagnetic Propagation and Geomagnetic Field Stability

P. J. COLEMAN, JR., AND C. P. SONETT

*National Aeronautics and Space Administration
Washington 25, D. C.*

The adjoining letter, Dessler [1961] argues that the interface between the solar wind and the geomagnetic field is stable and cannot, therefore, generate hydromagnetic waves. According to his interpretation, his argument may be stated as follows: He observes that, for the sudden commencement geomagnetic disturbances, the surface magnetograms show no marked change in the level of disturbances from the time before the sudden commencement until the initial phase of the storm. He uses the attenuation factors calculated by Francis and Karplus [1960] to the measurements of the 1 cps component of the magnetic disturbances obtained at the surface. From this observation, he concludes that, during the initial phase of these geomagnetic storms, the amplitude of hydromagnetic disturbances above the surface are less than 1γ . Assuming that the sudden commencement indicates the presence of the solar wind, he further concludes that the magnetic disturbances are not generated in the region between the geomagnetic field and the solar wind and that this interface is, therefore, inherently stable. He subsequently accounts for the hydromagnetic disturbances that are frequently observed at the earth as being produced by fluctuations in the density of the impinging solar wind. It is not our purpose to question whether the solar wind fluctuations generate any or not the hydromagnetic waves observed in the geomagnetic field. Rather, we wish to point out that the method by which Dessler uses surface measurements of magnetic activity to estimate conditions at great distances above the surface is not valid in the light of available data. Our data include observations obtained from the space probe, Pioneer I [Sonett, Smith, and Kelso, 1960; Sonett, Judge, Sims, and Kelso, 1960]. During the passage of the spacecraft

through the distant geomagnetic field, field strength fluctuations of large amplitude ($\Delta B/B \leq 10$) were detected.¹ Measurements obtained simultaneously with a surface magnetometer in the Borrego Desert (Campbell, private communication) indicated almost no surface geomagnetic activity. This magnetometer had a threshold of about 0.1γ in the frequency range of interest and is one of the instruments upon which Dessler relied for his data. Also, at the time that Pioneer I observed these disturbances, the A_p index was between 1 and 0. The predominant frequency component of these disturbances was about 0.1 cycle/sec. Frequencies of this value, according to the work of Francis and Karplus, should be less effectively attenuated than those at 1.0 cycle/sec. However, the observations indicate that even the lower frequency disturbances detected in the distant geomagnetic field were not measurable at the surface. Thus, Dessler's assumption that any large-amplitude hydromagnetic disturbances generated at the interface between the geomagnetic field and the solar wind would have been detected by the instruments which he mentions is not consistent with experiment.

Because of the apparent inconsistency of available data with Dessler's comments, we would like to consider briefly the problems that might be encountered in any treatment of hydromagnetic wave propagation in the exosphere. For example, the details of energy transport by such waves have not been established. The dispersionless character of very low frequency Alfvén waves suggests that, when the energy flux is constant, the amplitude of waves traveling inward through the geomagnetic field should decrease until the waves reach the region

¹ Large amplitude waves were observed also in the interface region on Pioneer V [Coleman, Sonett, Judge, and Smith, 1960].

in which the dependence of their velocity upon the ion density is of overriding importance compared with dependence on the field strength. However, the situation is probably complicated by the anisotropic behavior of the extraordinary mode and by the inhomogeneous nature of the medium to both the ordinary and extraordinary rays. Another problem arises when one attempts to ascertain the manner in which these disturbances, which, from a consideration of their amplitudes, structure, and velocities, appear to be shock-like phenomena in the distant field, are transformed into well-behaved waves as they enter the stronger fields nearer the surface. Further, we know of no treatment of the propagation of nonplanar hydromagnetic waves in the exosphere or of standing wave phenomena.

In summary, hydromagnetic disturbances as great as 100γ have been observed in the distant geomagnetic field with no associated effects observed by instruments of the type discussed by Dessler. The complexity of the observed phenomena makes it difficult to establish whether the waves are generated by instabilities at the interface or by fluctuations in the

intensity of the solar wind. However, the little empirical evidence at hand seems to indicate that it is dangerous to infer too much about disturbances of the type under discussion in the distant geomagnetic field on the basis of available ground observations.

REFERENCES

- Coleman, P. J., Jr., C. P. Sonett, D. L. Judge, and E. J. Smith, Some preliminary results of the Pioneer V magnetometer experiment, *J. Geophys. Research*, **65**, 1856-1857, 1960.
- Dessler, A. J., The stability of the interface between the solar wind and the geomagnetic field, *J. Geophys. Research*, **66**, 3587-3590, 1961.
- Francis, W. E., and R. Karplus, Hydromagnetic waves in the ionosphere, *J. Geophys. Research*, **65**, 3593-3600, 1960.
- Sonett, C. P., E. J. Smith, and A. R. Sims, Surveys of the distant geomagnetic field: Pioneer I and Explorer VI, *Space Research*, edited by H. Kallman-Bijl, North Holland Publishing Company, Amsterdam, pp. 921-937, 1960.
- Sonett, C. P., D. L. Judge, A. R. Sims, and J. Kelso, A radial rocket survey of the distant geomagnetic field, *J. Geophys. Research*, **65**, 55-68, 1960.

The Lifetime of Radiation Belt Protons with Energies between 1 Kev and 1 Mev

HAROLD LIEMOIHIN

*Geo-Astrophysics Laboratory, Boeing Scientific Research Laboratories
Seattle, Washington*

The lifetime of protons in the Van Allen radiation belts is limited by several loss mechanisms. During geomagnetically quiet periods the principal loss mechanisms are probably Coulomb scattering and charge exchange with the components of the partially ionized exosphere. The Coulomb scattering lifetime has been calculated as a function of proton energy and equatorial scattering density [Wentworth, MacDonald, and Singer [1959]]. The charge exchange lifetime is calculated here using experimentally measured cross sections. In using specific density models for the neutral and charged components of the exosphere, the two mechanisms are compared for proton energies between 1 kev and 1 Mev where charge exchange is important.

The radiation belt proton component at kev energies has not been studied experimentally owing to limitations of detector sensitivity. It has, however, received some theoretical consideration. Stuart [1959] has pointed out that the lifetime of less than a day prevents any appreciable density at kev energies unless a source is available. Recently, Jastrow [1960] has suggested that the outer zone of radiation consists primarily of protons from solar storms with energies near 10 kev. The observed electron component is then a consequence of the partitioning of energy with some of the thermal protons in the background exosphere.

Additional evidence for the existence of an appreciable kev proton component may be found in the primary auroral flux. Van Allen, McIlwain, and Ludwig [1959] have suggested that the flux emanates from the tip of the outer zone. Chamberlain [1957] has estimated the auroral proton flux to be 10^7 – 10^8 protons/cm² sec with energies in the kev range during geomagnetically quiet periods. Such a continuous flux of protons in the outer zone would require an appreciable steady state density.

The exosphere is believed to consist of thermal atomic hydrogen [Chapman, 1957] and protons and electrons [Storey, 1953] constrained to maintain electrical neutrality. The hydrogen density model obtained by Johnson and Fish [1960] is used here to calculate the charge exchange lifetime. The Coulomb scattering lifetime is calculated using the proton and electron density model assumed by Wentworth, MacDonald, and Singer [1959]. At altitudes from 3 to 5 earth radii, this latter model agrees closely with recent rocket measurements by Gringaus, Kurt, Moroz, and Shklovskii [1960] but is somewhat higher than observed at lower altitudes. The density models are very nearly identical in shape ranging from about 10^4 particles/cm³ at 1.5 earth radii to 20 at 5.5 earth radii.

The motion of nonrelativistic charged particles that are trapped in the dipolar geomagnetic field is well known [for example, see Welch and Whitaker, 1959; Wentworth, MacDonald, and Singer, 1959]. In this application it is sufficient to consider only the spiral motion between conjugate mirror points. The orbit is conveniently specified by its geocentric altitude at the geomagnetic equator r_0 and its mirror latitude λ_m .

A proton will generally make many traversals between its mirror points before being lost. Consequently, its calculated lifetime is not altered appreciably if the true instantaneous scattering density n is replaced by an average value over the orbit

$$\bar{n}(r_0, \lambda_m) = \int n ds / \int ds \quad (1)$$

where ds is an element of orbit arc length. The lifetime of a proton that traverses an appreciable variation in density can be easily evaluated with this simplification.

A convenient approximation for \bar{n} that is valid for both models of the charged and neutral

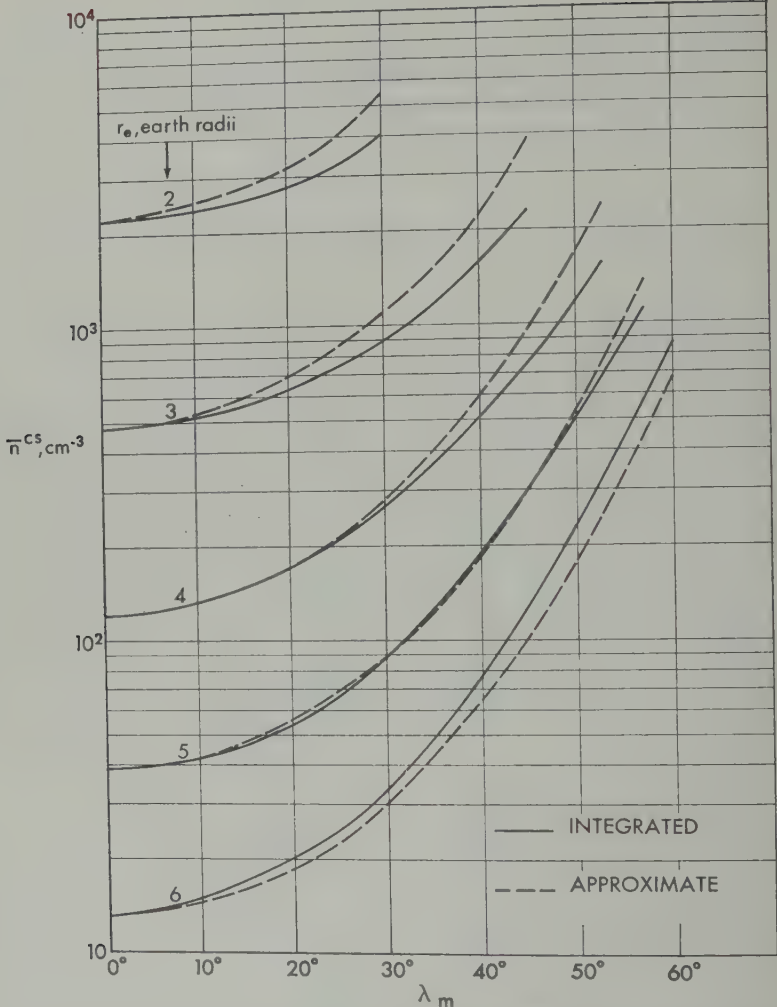


Fig. 1. Average Coulomb scattering density. Integrated and approximate values for the density \bar{n}^{CS} are plotted as a function of mirror latitude λ_m for several geocentric equatorial altitudes r_e .

components in the range $2 \leq r_e \leq 6$ earth radii is given by

$$\bar{n} = n_e(r_e) \sec^6 \lambda_m \tag{2}$$

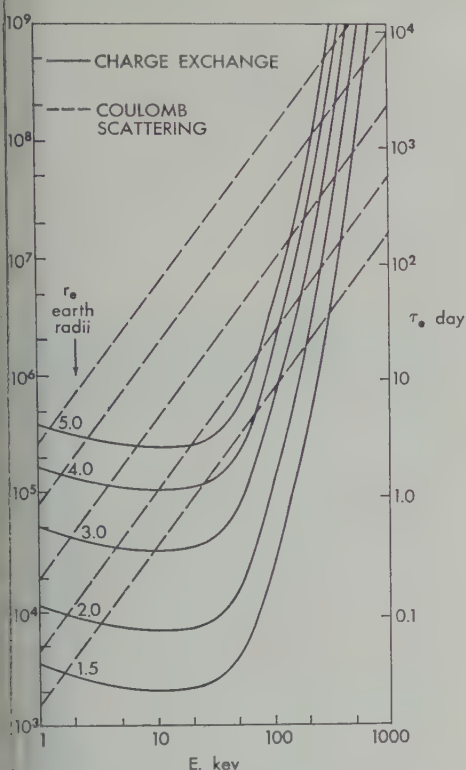
where n_e is the equatorial density. The integrated and approximate expressions for the average Coulomb scattering density \bar{n}^{CS} in a dipole field are plotted in Figure 1 as functions of λ_m for several r_e .

Since charge exchange involves a single interaction between a proton and a hydrogen atom, the lifetime for a proton that mirrors at the equator is given by

$$\tau_e^{CE} = (\sigma n_e^{(H)} v)^{-1} \tag{3}$$

where σ is the charge exchange cross section, $n_e^{(H)}$ is the equatorial density of atomic hydrogen, and v is the speed of the proton.

Experimental values of σ have recently been measured by Fite, Stebbings, Hummer, and Brackmann [1960] for proton energies from 0 to 40 keV. At higher energies σ is closely approximated by half of the cross section for protons and molecular hydrogen which has been extensively tabulated by Allison [1958] in a review paper. This approximation has been investigated theoretically by Tuan and Gerjuoy [1960], who found that the molecular value per atom and the atomic value agree closely for proton energies



2. Charge exchange and Coulomb scattering lifetimes. The equatorial lifetime τ_e is plotted as a function of proton energy E for several altitudes r_e .

400 keV but differ by 20 to 40 per cent at this energy. In the present application, the error above 400 keV does not affect the conclusions so that half of the molecular cross section can be used to approximate the atomic one below 400 keV.

Using the atomic hydrogen density model for the equatorial charge exchange lifetime may be plotted as a function of proton energy E and geocentric equatorial altitude r_e as shown in Figure 2. The equatorial Coulomb scattering lifetime τ_e^{cs} is also plotted in Figure 2 to illustrate the dominance of charge exchange at energies from 2 to 200 keV.

For radiation belt protons that mirror at the equator ($r_e \sim 3$ to 4 earth radii) have charge exchange lifetimes of about 1 day for energies from 2 to 75 keV. Above 75 keV the lifetime increases rapidly with energy reaching the order of a year at 200 keV.

For the lifetime for both mechanisms is in-

versely proportional to the average scattering density, the lifetime of a proton that has a mirror latitude λ_m is given approximately by

$$\tau = \tau_e n_e / \bar{n} = \tau_e \cos^6 \lambda_m \quad (4)$$

using equation 2. Hence, the lifetime is shortened appreciably for high mirror latitudes. For example, a proton that mirrors near the tip of the outer belt, 1 earth radius above the earth's surface, has a lifetime about 0.15 as long as at the equator.

REFERENCES

- Allison, S. K., Experimental results on charge-exchanging collision of hydrogen and helium atoms and ions at kinetic energies above 0.2 keV, *Revs. Modern Phys.*, **30**, 1137-1168, 1958.
- Chamberlain, J. W., On a possible velocity dispersion of auroral protons, *Astrophys. J.*, **126**, 245-252, 1957.
- Chapman, S., Speculations on the atomic hydrogen and the thermal economy of the upper atmosphere, in *The Threshold of Space*, edited by M. Zelikoff, Pergamon Press, London, pp. 65-72, 1957.
- Fite, W. L., R. F. Stebbings, D. G. Hummer, and R. T. Brachmann, Ionization and charge transfer in proton-hydrogen atom collisions, *Phys. Rev.*, **119**, 663-668, 1960.
- Gringaus, K. I., V. G. Kurt, V. I. Moroz, and I. S. Shklovskii, Ionized gas and fast electrons in the vicinity of the earth and in interplanetary space, *Doklady Acad. Nauk SSSR*, **132**, 1062-1065, 1960.
- Jastrow, R., Geophysical effects of the trapped particle layer, *Revs. Modern Phys.*, **32**, 947-950, 1960.
- Johnson, F. S., and R. A. Fish, The telluric hydrogen corona, *Astrophys. J.*, **131**, 502-515, 1960.
- Storey, L. R. O., An investigation of whistling atmospherics, *Phil. Trans. Roy. Soc. London, Ser. A*, **246**, 113-141, 1953.
- Stuart, G. W., Satellite-measured radiation, *Phys. Rev. Letters*, **2**, 417-418, 1959.
- Tuan, T. F., and E. Gerjuoy, Charge transfer in molecular hydrogen, *Phys. Rev.*, **117**, 756-763, 1960.
- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1958, *J. Geophys. Research*, **64**, 271-286, 1959.
- Welch, J. A., Jr., and W. A. Whitaker, Theory of geomagnetically trapped electrons from an artificial source, *J. Geophys. Research*, **64**, 909-922, 1959.
- Wentworth, R. C., W. M. MacDonald, and S. F. Singer, Lifetimes of trapped radiation belt particles determined by Coulomb scattering, *Phys. Fluids*, **2**, 499-509, 1959.

(Received March 21, 1961;
revised July 25, 1961.)

Discussion of the Letter by R. A. Hanel, 'Determination of Cloud Altitude from a Satellite'

G. YAMAMOTO AND D. Q. WARK

U. S. Weather Bureau, Washington, D. C.

Recently Hanel [1961] made the interesting suggestion that the estimation of cloud-top altitude is possible by measuring from a satellite the absorption of reflected solar radiation by a CO_2 band. We agree with this suggestion in principle, but at the same time we consider that the reflected ray from the cloud top must be strong and that its absorption by the medium above the cloud must be considerable. The

choice of the absorption band is therefore very important for this problem. Hanel proposed the use of the $2\ \mu$ CO_2 band. The $0.76\ \mu$ O_2 band seems to be preferable, for the following reasons:

1. While the $2\ \mu$ CO_2 band is overlapped by the wing of the $1.87\ \mu$ H_2O band, which makes the analysis of the observation somewhat difficult, there is no overlapping by other bands in the $0.76\ \mu$ O_2 band, as is shown in Figure 1.

2. The absorption of solar radiation by the $0.76\ \mu$ O_2 band is somewhat greater than that by the $2\ \mu$ CO_2 band, giving better contrast with the adjacent reference intervals.

3. The reflectivity of clouds at around $0.76\ \mu$ seems to be greater than that at $2\ \mu$, and the solar intensity is greater.

4. At $0.76\ \mu$ more sensitive detectors are available than at $2.0\ \mu$, and better signal-to-noise ratios are possible.

Two aspects of this method require further investigation. First, the absorption characteristics of the $0.76\ \mu$ band are not yet known to the desired accuracy, but we hope to study them in the near future. Second, clouds may not be regarded as simple diffuse reflectors, but rather the absorption along the scattering paths within clouds must be considered.

Acknowledgment. This work was supported by the National Aeronautics and Space Administration.

REFERENCES

- Hanel, R. A. Determination of cloud altitude from a satellite, *J. Geophys. Research*, **66**, 1300, 1961.
- Howard, J. N., D. L. Burch, and D. Williams. Near-infrared transmission through synthetic atmospheres, *Geophys. Research Papers*, no. 46, AFCRC-TR-55-213, 1955.
- Langley, S. P., and C. G. Abbot, *Ann. Astrophys. Observatory Smithsonian Inst.*, **1**, 1900.

(Received July 20, 1961.)

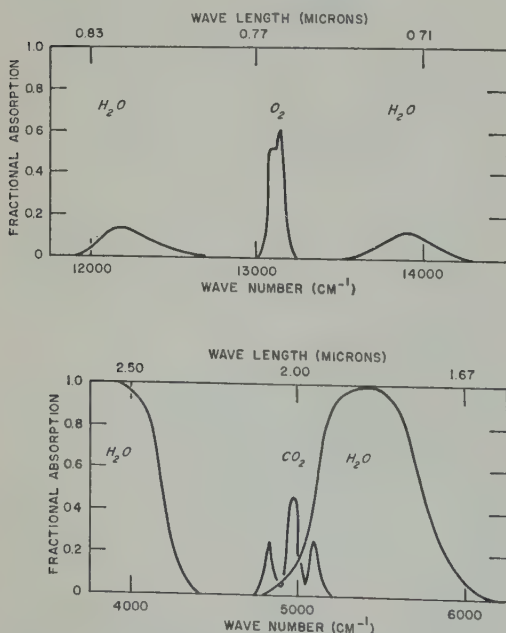


Fig. 1. The absorption of solar radiation at the earth's surface, expressed in fractional absorption, by the H_2O , CO_2 , and O_2 bands around 0.76 and $2\ \mu$ in average October, 30° – 40°N atmospheric conditions at normal incidence ($2.62\ \text{cm pr. water}$). The absorption by H_2O and CO_2 bands is based on data of Howard, Burch and Williams [1955] and the O_2 band absorption on those of Langley and Abbot [1900].

Evaluation of the Special World Interval Program during the IGY

MARTIN E. NASON

*CRPL, National Bureau of Standards
Boulder, Colorado*

is of interest to review the degree of success was achieved in declaring Special World Intervals (SWI) during the International Geophysical Year. These declarations were made at each day by the IGY World Warning Agency operated by the North Atlantic Radio Warning Service of the National Bureau of Standards, at Fort Belvoir, Virginia. It will be noted that, during IGY, the objectives of the Special World Interval program were not the same as the usual forecasts of geomagnetic disturbances. Special World Intervals were to be issued only when there was good probability of a significant geomagnetic disturbance would occur or continue during the 24-hour period following the declaration. The SWI program was developed so that special and concentrated observations could be coordinated among the

various scientific disciplines during the prestorm, start, and more severe phases of magnetic storms.

In Table 1 are listed the dates during the IGY on which a Special World Interval was in progress and the dates on which A_p , a daily measure of geomagnetic activity, was equal to or greater than 30 (a value generally agreed upon as being indicative of significant disturbance). Shown also are the dates common to the SWI and A_p columns.

The actual distribution of storms and Special World Intervals is shown in Table 2. During the 18-month period of the IGY, Special World Intervals were declared or in progress on 45 days. A significant geomagnetic disturbance either began or was in progress on 25 of these days. Neglecting persistence for the moment, and using the chi-square test, the probability of

TABLE 1

Year/Month	Dates of SWI	Dates $A_p \geq 30$	Common Dates
1957 July	1, 2, 3	1, 2, 3, 5	1, 2, 3
Aug.	24, 29, 30	6, 13, 30, 31	30
Sept.	2, 3, 4, 12, 13, 14	2, 3, 4, 5, 6, 13, 14, 21, 22, 23, 24, 29, 30	2, 3, 4, 13, 14
Oct.	22, 23	14	...
Nov.	26, 27 (in progress)	7, 25, 26, 27	26, 27
Dec.	...	11, 31	...
1958 Jan.	...	1, 18, 21	...
Feb.	...	5, 6, 11, 12, 17, 18, 21	...
Mar.	5, 15, 23, 24, 25, 30, 31	5, 6, 12, 13, 18, 19, 21, 25, 30	5, 25, 30
Apr.	...	2, 4, 16, 17, 18, 29	...
May	...	14, 26, 29, 31	...
June	6, 7, 8, 20, 21, 22	1, 7, 9, 10, 21, 22, 28, 29	7, 21, 22
July	8, 9, 30, 31	8, 9, 18, 21	8, 9
Aug.	17, 18, 24, 27, 28	17, 22, 24, 27	17, 24, 27
Sept.	...	3, 4, 5, 16, 25	...
Oct.	23, 24, 25 (in progress)	22, 23, 24, 27, 28	23, 24
Nov.	26, 27
Dec.	13, 14	4, 13, 17, 18	13

Total number of dates

45

88

25

TABLE 2

SWI	Storm			
	Yes	Yes	No	
		25	20	
SWI	No	63	441	504
		88	461	549

obtaining a distribution as good as or better than that shown in Table 2 by chance is only 2×10^{-13} . For comparison, Table 3 shows the distribution that would be expected from chance. If one becomes more severe in evaluation and does not allow credit for the four SWI that were issued after a disturbance had begun, one still finds that the probability of obtaining a distribution giving 21 successes is about 5×10^{-7} .

The last half of September 1957 produced several failures in the above study, since, although the IGY World Warning Agency was aware of pending major disturbance, it did not issue any SWI. This decision was based on the fact that six SWI had already been declared during the first half of that month, four of them covering periods of disturbance.

Since magnetic activity each day is not independent of the activity on the preceding day, nor is the declaration of an SWI entirely independent of the decision reached the day before, the effects of persistence should be taken into account. Since the average duration of a geomagnetic disturbance during the IGY is 1.6

TABLE 3

SWI	Storm			
	Yes	Yes	No	
		7	38	
SWI	No	81	423	504
		88	461	549

days and the average length of a SWI was 1.6 days during the same period, it has been judged that the number of independent observations limited to about $549/2$, or 274.5. There were 21 storms that began on the day a Special World Interval began and 40 that began on days with no SWI in effect. Using these cell frequencies as the number of independent observations reasonable to have been possible during IGY, the contingency table shown in Table 4 is constructed. The probability of obtaining a distribution as good as or better than that shown by Table 4 is about 5×10^{-7} .

TABLE 4

SWI Start	Storm Start			
	Yes	Yes	No	
		13	8	21
SWI Start	No	40	213.5	253.5
		53	221.5	274.5

From these studies it seems clear that the forecasting techniques used by the IGY World Warning Agency were certainly valid and that the objectives of the Special World Interval program were achieved. This is particularly true when we remember that the Special World Interval was to cover a period of time from a few hours before the start of a geomagnetic disturbance into the maximum phase of the disturbance.

REFERENCES

Annals of IGY, 7, 12-14, Pergamon Press, 1959.
Crow, Edwin L., *J. Am. Statistical Assoc.*, 47, 303-304, 1952.
Eisenhart, C., M. W. Hastay, W. Allen Wallis, *Techniques of Statistical Analysis*, chapter 7, 1946.
International Scientific Radio Union, *Informations Bull.* 88, 38-44, 1954.

(Received July 13, 1961.)

Magnetic Field Micropulsations and Electron Bremsstrahlung

WALLACE H. CAMPBELL

Central Radio Propagation Laboratory, National Bureau of Standards
Boulder, Colorado

note is to report the observation of
ed magnetic micropulsation activity in
roral zone near College, Alaska, simul-
s with increases in the intensity of
strahlung from energetic electrons. Meas-
nts of the magnetic field fluctuations were
with a 2-meter-diameter loop antenna of
turns with its axis in the magnetic north-
direction. The system had a flat response

to magnetic flux density in the frequency range
of 0.4 to 0.04 c/s. Balloon flights were made near
the same site [Brown, 1961] in June and July
1960. Using Geiger counters, Brown observed
bremsstrahlung from bombarding electrons
having energies greater than 50 kev. Figure 1
shows the X-ray bursts, representing high-
energy electron influx, coincident with the mag-
netic field micropulsation amplitudes. The balloon

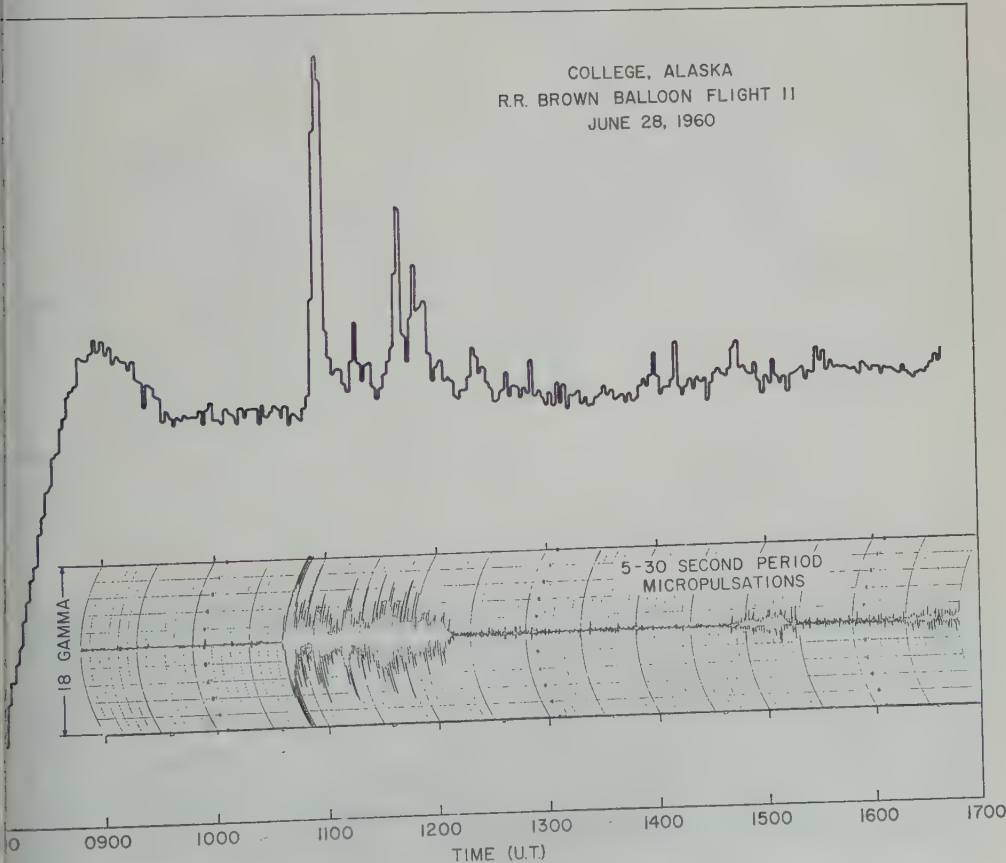


Fig. 1. Sudden commencement of micropulsation activity coincident with auroral electron
bremsstrahlung near College, Alaska, on June 28, 1960. (The Pfitzer maximum at 0855 UT is a
balloon-ascent phenomenon.)

reached its floating altitude at about 0930 UT; note the Pfozter maximum during this ascent and a gradual background count increase as the balloon loses altitude during the flight. The large burst at 1050 UT and subsequent activity are shown on both data samples.

Balloon measurements of electron bremsstrahlung using rapid time response scintillation counters (cf. Fig. 11 of *Anderson and Enemark* [1960]) give some indication of the existence of bunching in the incoming electron density.

Further observations should show whether such short-term fluctuations are also simultaneous with the magnetic field pulsations.

REFERENCES

- Brown, R. R., Balloon observations of auroral zone X rays, *J. Geophys. Research*, **66**, 1379-1388, 1961.
Anderson, K. A., and D. C. Enemark, Balloon observations of X rays in the auroral zone, *J. Geophys. Research*, **65**, 3521-3538, 1960.

(Received July 10, 1961.)

Short-Term Phase Perturbations Observed at 18 kc/s

C. F. SECHRIST AND K. D. FELPERIN

*HRB-Singer, Inc.
State College, Pennsylvania*

purpose of this letter is to report on the observation at State College of small, short-term fluctuations, on the 18-kc/s signal transmitted from NBA, Balboa, Panama Canal Zone, north-south propagation path length of 3500

feet. The transmitter frequency is highly stable. The NBA carrier frequency of 18 kc/s is kept constant to about ± 1 part in 10^{10} by means of a highly stable quartz-crystal oscillator, and the oscillator is regulated in accordance with the time (by means of the clocks of the cesium type), and the phase is controlled in State College, Pennsylvania, by multiplying the signal to 2 kc/s by means of a frequency standard (1 part in 10^9 per second). The 2-kc/s signal is 'squared,' and used to modulate the intensity of a Tektronix 515A oscilloscope, whose time base is triggered by a signal derived from the frequency standard. A horizontal line trace appears on the oscilloscope screen, whose end point moves in proportion to the phase of the NBA signal. This display is highly sensitive, corresponding to 2.8° per centimeter of deflection, and measurements to be made to 0.5° . According to the waveguide mode theory of Wait [1959], a phase change of 2.8° corresponds to a change of ionospheric height of 50 meters. This method of observation thus provides a powerful tool for

investigating small changes in the lower ionosphere which are usually not observed by other methods.

Apart from the usual sunrise and sunset changes of phase, our records show quasi-sinusoidal phase changes with periods in the range 10 to 20 seconds and amplitudes of 2° to 6° . These have been observed in the daytime but not during the night, making it extremely unlikely that they are produced by instrumental variations.

A 2-minute sample record is shown in Figure 1 for the period 1446 to 1448 EST on March 9, 1961. The separate lines correspond to single NBA pulses. The quasi-sinusoidal oscillations between 1446 and 1447 have periods of about 13 seconds and peak-to-peak amplitudes of about 3° . On the waveguide mode theory of Wait [1959], this corresponds to a height change of about 300 meters of the upper part of the ionosphere-earth waveguide, here assumed to be uniform over the propagation path. Other interpretations are possible, of course, for example, localized disturbances in the ionosphere of larger magnitudes. As a check on the instrumentation, the diurnal phase change corresponded to a height change of 15 km, in good agreement with the experimental results of other workers [Pierce, 1955; Crombie, Allen, and Penton, 1956; Crombie, Allen,

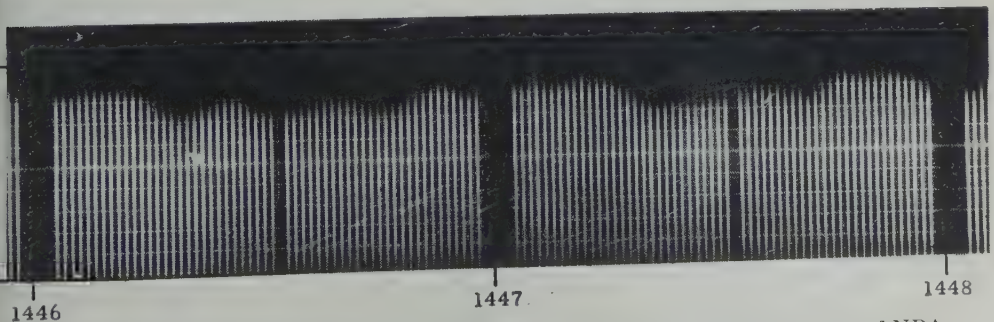


Fig. 1. Sample record for March 9, 1961, 1446 to 1448 hours EST. Pulse repetition rate of NBA is 1 pps, and pulse width is 0.3 seconds. S/N ratio approximately 20 db.

and Newman, 1958; Bain, Bracewell, Straker, and Westcott, 1952].

We are at present investigating the correlation of the small 18-kc/s phase fluctuations with geomagnetic micropulsations which have comparable quasi-periods and an average diurnal behavior that exhibits greatest activity in the daytime [Campbell, 1959].

REFERENCES

- Bain, W. C., R. N. Bracewell, T. W. Straker, and C. H. Westcott, The ionospheric propagation of radio waves of frequency 16 kc/s over distances of about 540 km, *Proc. IEE*, 99, 250-260, 1952.
- Campbell, W. H., Studies of magnetic field micropulsations with periods of 5 to 30 seconds, *J. Geophys. Research*, 64, 1819-1826, 1959.
- Crombie, D. D., A. H. Allen, and M. Newnham, Phase variations of 16 kc/s transmissions from Rugby as received in New Zealand, *Proc. IEE B*, 301-304, 1958.
- Crombie, D. D., A. H. Allen, and W. A. Penrice, Frequency variations in New Zealand of 16 kc/s transmissions from GBR Rugby, *Nature*, 178-179, 1956.
- Pierce, J. A., The diurnal carrier phase variations of a 16 kc transatlantic signal, *Proc. IRE*, 43, 584-588, 1955.
- Wait, J. R., Diurnal change of ionospheric height deduced from phase velocity measurements at VLF, *Proc. IRE*, 47, 998, 1959.

(Received June 22, 1961.)

Comments on K. Vozoff's Paper 'Calibration of Pulsation Detector Coils'

JAMES R. WAIT

*National Bureau of Standards
Boulder, Colo.*

Just-published communication, Vozoff discusses the validity of the free-space solution in calibrating detector coils for pulsation work. In connection with this solution, Vozoff says that he was not able to find a solution in the literature for the magnetic field at the center of a large circular loop lying over a conducting half-space. Actually, all one needs to do is to look at the mutual impedance of two coaxial loops, one of which is of a very small radius and the other of radius a . This is easily obtained by multiplying the E_ϕ component of a magnetic dipole of moment $2\pi a$ and dividing this, in turn, by the current I in the magnetic dipole. Vozoff's ratio $A - iB$ is the ratio of E_ϕ of a magnetic dipole on the conducting surface to the corresponding value of E_ϕ if the magnetic dipole were in free space. The ratio is given explicitly in an earlier paper of Wait [Wait, 1951] as equation 33.

Wait only refers to a later paper of mine [Wait, 1954], where I had given a general expression for the fields of a circular loop on a conducting surface. He says the field appears to 'blow up' at the center of the loop. This is not so if one goes through the indicated differentiations correctly before setting $\rho = 0$. Alternatively, one could retrieve Vozoff's integral expression for this by simply making use of the series formula for the Bessel function J_0 . However, this special case is obtained readily by using the method indicated in the preceding paragraph.

An important limitation, not explicitly mentioned by Vozoff, is the idealization of a constant current in the large circular loop. If the loop is actually lying on the ground, there may be a significant decrease of the current as one moves, along the circumference, away from the terminals. It is recommended, for work of this kind, that the calibrating loop be raised above the ground to a height of a meter or so. This will tend to minimize the nonuniformity of the loop current.

The inhomogeneity of the ground is another factor that could be taken into account in problems of this kind. Unfortunately, the integrals for the stratified half-space are quite a bit more complicated [Wait, 1951]. Recently, however, the various integral formulas [Wait, 1958] for a two-layer ground, with arbitrary dipole excitation, have been evaluated by the U. S. Geological Survey computing staff.

REFERENCES

- Vozoff, K., Calibration of pulsation detection coils, *J. Geophys. Research*, 66(6), 1983-1984, 1961.
- Wait, J. R., The magnetic dipole over the horizontally stratified earth, *Can. J. Phys.*, 29, 577-592, 1951.
- Wait, J. R., Mutual coupling of loops lying on the ground, *Geophysics*, 19, 290-296, 1954.
- Wait, J. R., Induction by an oscillating magnetic dipole over a two-layer ground, *Appl. Sci. Research*, B, 7, 73-80, 1958.

(Received June 28, 1961.)

Note on the Autocorrelation Coefficient of K_p and Its Relation to M Regions

W. R. PIGGOTT

*Department of Scientific and Industrial Research, Radio Research Station
Ditton Park, Slough, Bucks, England*

In a valuable paper, Ward [1960] has obtained the variation of the autocorrelation coefficient for K_p over a period of 400 days. The period studied, April 1, 1951, to March 31, 1956, covers the period in the solar cycle in which M -region-type magnetic disturbances are common.

The significant unexplained features of Ward's results are: (a) the 27-day periodicity becomes less clear after about the third cycle but is found, with large amplitude, between the tenth and fifteenth cycles, i.e., after a delay of about 1 year; (b) the autocorrelation is high but shows no significant oscillations for delays centered on 6 months.

The purpose of this note is to point out that these variations are consistent with the theory of the existence of long-lived M regions on the sun and provide additional evidence that they have a long life and are concentrated near two bands of solar latitude, one in the northern and one in the southern hemisphere.

For the epoch in the solar cycle studied by Ward, the most active zone of latitudes in the northern hemisphere of the sun faced the earth at one equinox whereas the corresponding zone in the southern hemisphere faced the earth in the next equinox. It is therefore to be expected that terrestrial magnetic activity, measured by K_p , should be great at the two equinoxes and small at the two solstices. Ward's autocorrela-

tion function for K_p is, on the whole, greater at the equinoxes and smaller at the solstices in agreement with this expectation.

What is particularly interesting is that the larger values of the autocorrelation coefficient show a 27-day period in the first and third equinoctial periods, when one and the other active latitude band on the sun came opposite the earth, but not near the second equinox when the other active latitude band faced the earth. This fact is consistent with the hypothesis of the distribution of active and inactive areas on the sun, which is remarkably stable, so that when the same active area came opposite the earth after 1 year the 27-day repetition period was still noticeable. Ward's results, however, a different set of active areas, in the other solar hemisphere, came opposite the earth at the second equinox, although there was a general increase in K_p during this period. In the autocorrelation coefficient, there was no detailed relation to the other set, and no 27-day periodic relation to it.

The results indicate that the angles over which the solar particles were emitted were very small and were situated in both the meridian and the equatorial planes.

REFERENCE

- Ward, F. W., *J. Geophys. Research*, **65**, 2359-2364, 1960.

(Received April 29, 1961.)

Author's Comments on the Previous Discussion

FRED WARD

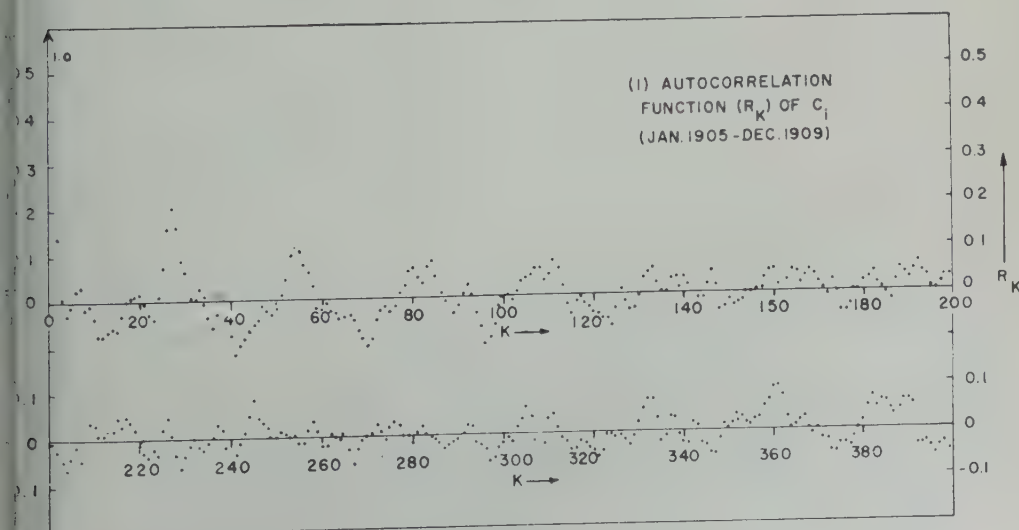
*Geophysics Research Directorate, Air Force Cambridge Research Laboratories
Bedford, Massachusetts*

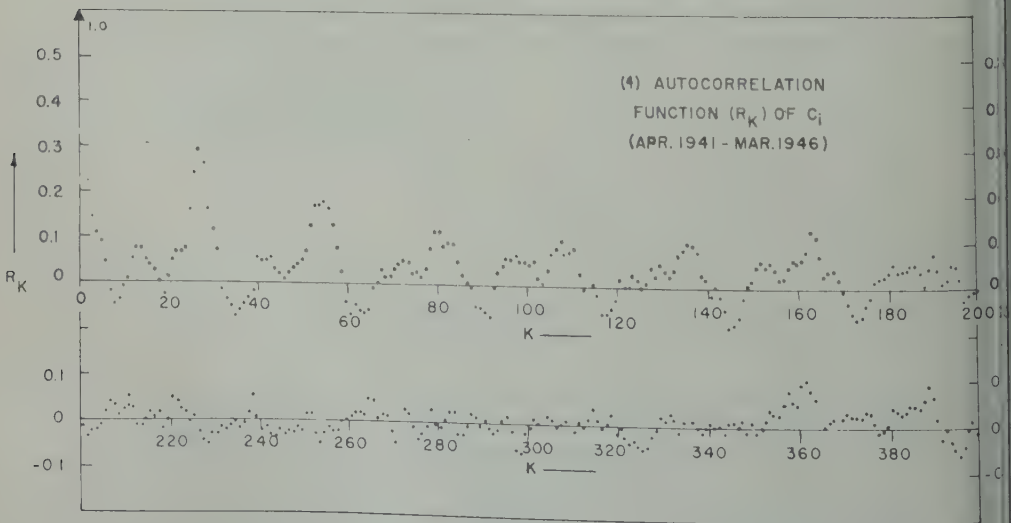
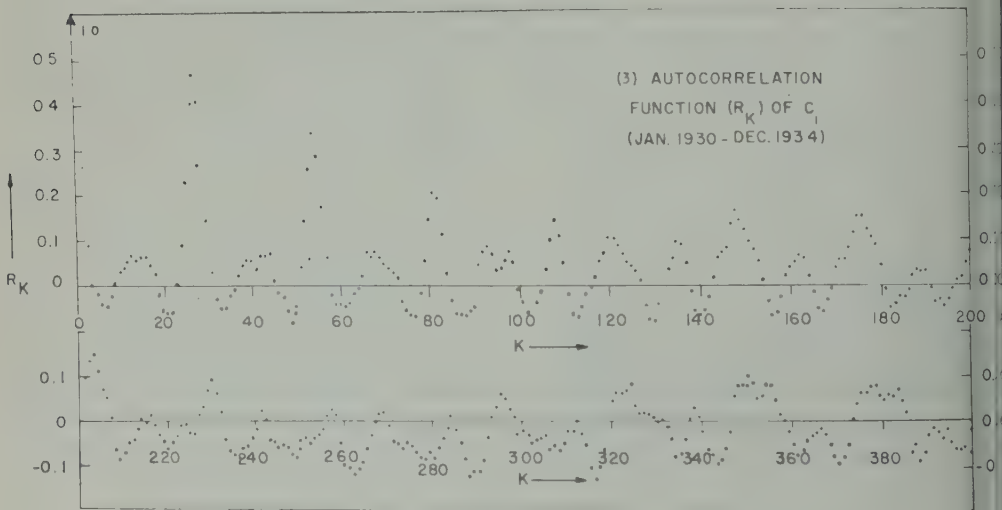
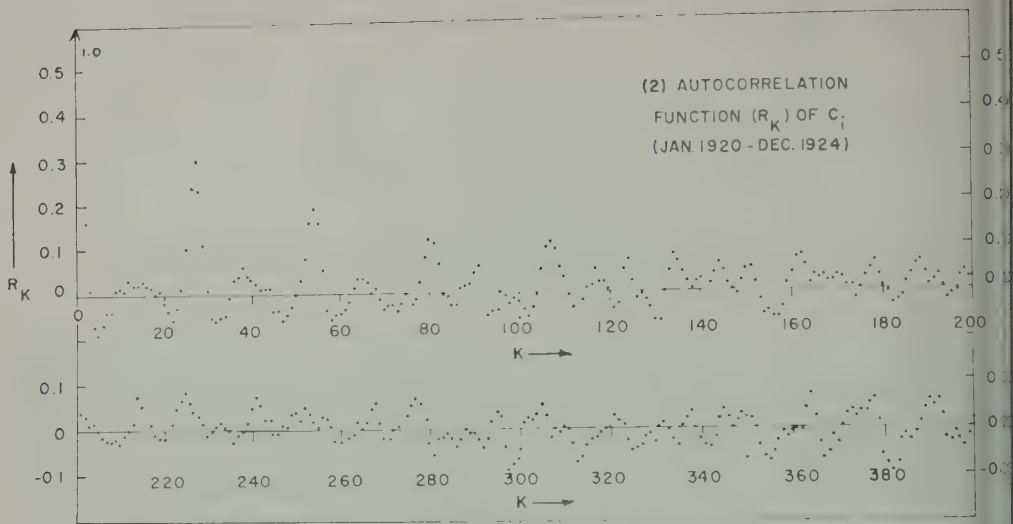
Dr. Piggott's discussion of Figure 2a of my paper is very interesting. Since one step in the computation of the spectra is the determination of the autocorrelation function, it was easy to compute the autocorrelation function for other periods of time.

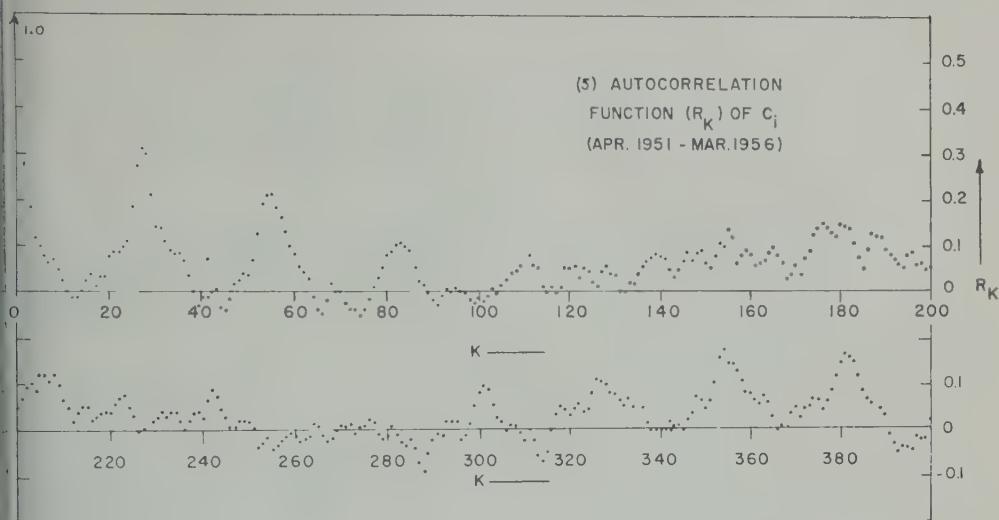
Figures 1-5 show the autocorrelation function R_K for four 5-year periods around sunspot maximum and one 5-year period around sunspot minimum. C_1 is used for comparison with past work. Figure 5 corresponds to the same period as Figure 2a in the original paper. Inspection of this figure and Figure 2a shows an expected close similarity. The most obvious feature of the five autocorrelation functions is their 27-day variation. In the period 1930-1934 (Figure 1) this variation is strong throughout the entire range of 400 lags. In the 1920-1924 period (Figure 2) the 27-day variation rapidly loses its identity with increasing lag. The behavior of the autocorrelation function in the 1941-1946 period (Figure 3) lies somewhere between these two extremes.

The period discussed in this correspondence (1951-1956, Fig. 5) is somewhat different from all three. The one period around sunspot maximum (1905-1909, Fig. 1) shows an autocorrelation function which again is somewhat different from all the others but at the same time has many similar characteristics. Each shows some peculiarities which make speculation tempting, but on the basis of the diversity of these results it is apparent that one cannot generalize about variations longer than three solar rotations from one autocorrelation function.

The autocorrelation function contains no phase information. That is, all periods of the year contribute equally to the autocorrelation at each lag. In Figure 5 there is a broad maximum centered around 180 days. This maximum cannot be equated to the well-known equinoctial maxima in geomagnetic activity. However, this approximate 6-month period is most likely the cause of the maximum in the autocorrelation function. Similarly, the recurrence of the 27-day







on at a lag of around 1 year cannot necessarily be interpreted as resulting from time-varying earth-sun geometry.

There does appear to be a slight enhancement of 27-day variation around 1 year in all five

autocorrelation functions. This is consistent with the idea of the existence of long-lived M regions on the sun.

(Received July 27, 1961.)

The Earth's Free Oscillations Observed on Earth Tide Instruments at Tiefenort, East Germany¹

W. BUCHHEIM

*Institute für theoretische Physik und Geophysik der Bergakademie Freiberg
Freiberg, East Germany*

S. W. SMITH

*Seismological Laboratory, California Institute of Technology
Pasadena, California*

The great Chilean earthquake of May 22, 1960, wrote unusual long-period records on horizontal pendulum seismographs designed for recording earth tides at Tiefenort, East Germany (50°52'N, 10°57'E). A section of the EW record is reproduced in Figure 1. It shows the foreshock at 1032 GCT (Mag 7½) and the main shock at 1911 GCT (Mag 8½) along with the semidiurnal earth tide.

Spectral analysis of the continuing long-period disturbance on this record showed a typical line spectrum characteristic of the earth (Fig. 2). The two horizontal components respond to both horizontal acceleration and tilt; thus there is no way of separating spheroidal from toroidal modes at this station except on the basis of frequency. For all but the lowest modes the spheroidal and toroidal modes are so close in frequency that even this criterion cannot be

used. Uncertainties in timing plus the fact only 22 hours of record could be digitized prevented good frequency resolution. In spite of these difficulties, the cross spectrum of the EW component and a comparable EW pendulum seismograph located at Pasadena (Fig. 3) has no doubt that the spectral peaks observed at Tiefenort do represent coherent vibrations of the entire earth.

The lowest mode recorded with any certainty is ${}_0S_4$ with a period of 25.8 minutes, and the highest mode is ${}_0S_{16}$ with a period of 6.78 minutes. Observed periods and tentative mode identifications are presented in Table 1, where the notations ${}_nS_l$ and ${}_nT_l$ indicate the n th radial overtone of surface harmonic dependence of order l . The identifications in Table 1 depend heavily on observations made with vertical instruments and reported by Ness, Harrison, and Smith [1961] and by Benioff, Press, and Smith [1961]. Although the periods presented here are

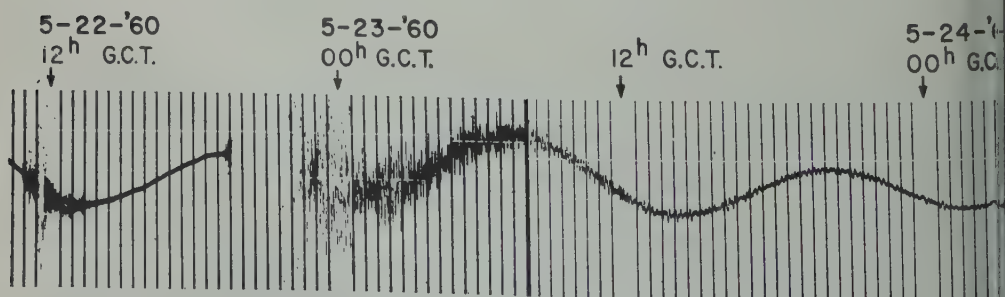


Fig. 1. EW component of the Tiefenort earth tide recording of the Chilean earthquake of May 22, 1960.

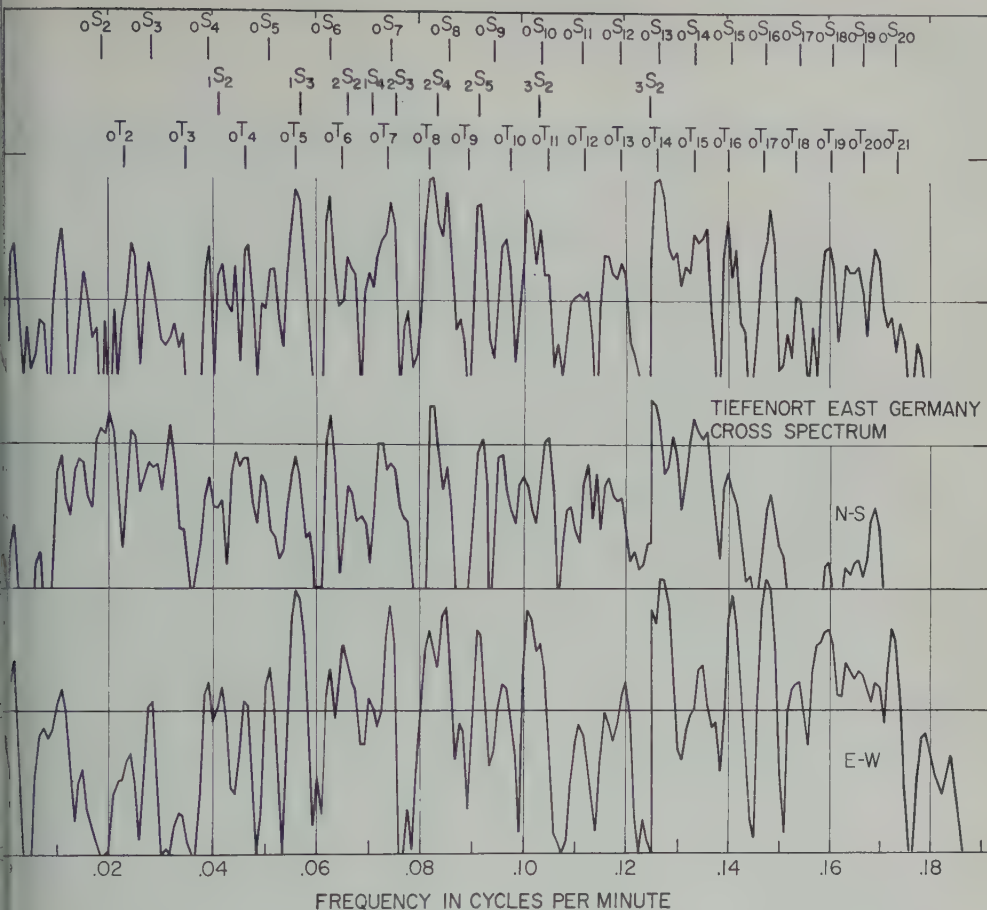
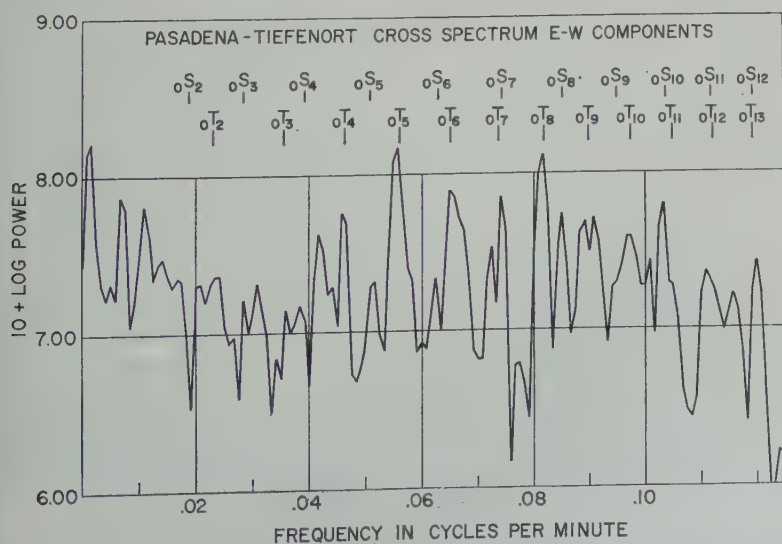


Fig. 2. Spectra and cross spectrum of NS and EW components at Tiefenort. Note that the vertical scale changes by a factor of 10 at 0.125 cpm. Mode designations are for a Gutenberg model earth.



3. Cross spectrum of EW components at Tiefenort and Pasadena. Mode designations are for a Gutenberg model earth.

TABLE 1. Free Oscillations Observed at Tiefenort

Period, min	Mode
25.8	${}_0S_4$
24.0	${}_1S_2$
21.6	${}_0T_4$
19.7	${}_0S_5$
17.9	${}_0T_5$ or ${}_1S_3$
15.3	${}_0T_6$
13.5	${}_0S_7$ or ${}_0T_7$
12.2	${}_0T_8$
11.8	${}_0S_8$
11.0	${}_0T_9$?
10.3	${}_0T_{10}$
9.93	?
9.68	${}_0T_{11}$
9.02	${}_0T_{12}$
8.63	?
8.33	${}_0S_{12}$ or ${}_0T_{13}$
7.87	${}_0S_{13}$
7.43	${}_0S_{14}$
7.10	${}_0S_{15}$
6.78	${}_0S_{16}$

precise than previously published values, and are at the present time the only available observations of the free oscillations in the eastern hemisphere. The phase differences and Q ratios provided by the cross spectral analysis of the two horizontal components at this station and of the EW components at Pasadena and Tiefenort have been used by Smith [1961] to infer some of the dynamic properties of the source.

REFERENCES

- Benioff, H., F. Press, and S. Smith, Excitation of the free oscillations of the earth by earthquakes, *J. Geophys. Research*, **66**(2), 605-619, 1961.
- Ness, N., J. Harrison, and L. Slichter, Observations of the free oscillations of the earth, *J. Geophys. Research*, **66**(2), 621-629, 1961.
- Smith, S. W., An investigation of the earth's free oscillations, Thesis, California Institute of Technology, 31 pp., 1961.

(Received July 17, 1961.)

Discussion of Paper by John F. Mink, 'Distribution Pattern of Rainfall in the Leeward Koolau Mountains, Oahu, Hawaii'

G. REYNOLDS

*North of Scotland Hydro-Electric Board
Edinburgh, Scotland*

er to Mink's [1960] findings and Oelsner Basso's [1961] conclusion that rainfall increases with altitude after a certain level of exposure has been reached. This is totally in accordance with all British experience which shows that precipitation increases with altitude up to a limit, but is readily explicable when viewed in the light of the different siting and exposure of rain gages officially recommended by the U. S. and British authorities. In the U. S. rain gages are installed so that their exposure is 2 inches above ground and if there are doubts for believing that the site is windy, as is often the case in mountainous districts, the gage is placed in a natural hollow in the ground to protect it from strong wind causing excessive turbulence around it. If a suitable hollow is not available the gage is mounted within a 10-foot-tall earthen or turf wall with a 1 in 4 slope on the windward side and a crest in the plane of the gage rim. In the U. S. this is not claimed to eliminate all effects of wind exposure.

In the U. K. rain gages on the other hand are mounted so that the rim is 31 inches above ground. Tests conducted as long ago as the last century showed that the results in the gage catching much less rain at 12 inches and that this loss increases markedly with wind speed. The differences noted by Kurtyka [1953] in his Table 6 are particularly marked at relatively sheltered sites. In exposed conditions the loss would be many times greater. British rain catch experiments [Braak, 1945] have demonstrated that the Nipher shield often gives false readings. U. S. gages to mitigate this nuisance are of little use.

Mink's finding that rainfall increases with altitude in the valley bottom (his gulch gages) is reasonable, but varying exposure to wind has

falsified all his ridge comparisons. Rainfall is increasing with height but so is exposure to wind. There comes a point, somewhere between gages 10 and 11, where the second effect assumes control and produces the reduced catch as observed. Even before this, the loss due to wind exposure has been increasing with altitude and the observed gage differences plotted in his Figure 2 are not true rainfall differences.

Lest the terms of this note appear smug it is only fair to point out that the U. S. gage procedure has been devised to reduce the likelihood of gages becoming buried in snow drifts, a problem which British meteorologists have rarely to face. The writer contends that neither gage is perfect.

The American gage is better for snow catch, for which purpose the British procedure is poor, especially if the gage is not tended daily (snow constitutes only a small percentage of our total precipitation), but for rainfall measurement the British gage is almost perfect. At least it is fully representative under all conditions, while the American exposure renders comparative work impossible if the gage sites are different in their wind characteristics.

Oelsner and Basso's Chilean catches are subject to the same exposure errors.

REFERENCES

- Braak, C., Influence of wind on rainfall measurements, *Mededeel. en Verhand. Koninkl. Ned. Meteorol. Inst.*, no. 48, 1945.
- Kurtyka, J. C., *Precipitation Measurement Study*, State of Illinois Water Survey Div., p. 8, 1953.
- Mink, J. F., Distribution pattern of rainfall in the leeward Koolau Mountains, Oahu, Hawaii, *J. Geophys. Research*, 65, 2869-2876, 1960.
- Oelsner, J. F., and E. S. Basso, Discussion on the above, *J. Geophys. Research*, 66, 1998-2000, 1961.

Author's Reply to the Preceding Discussion

JOHN F. MINK

*Board of Water Supply
Honolulu, Hawaii*

Arguments essentially the same as Reynolds' about the effect of wind on rain catch were, of course, carefully considered in preparation of my paper and were critically reviewed by others. Such discussions, however, can become quite prolix, even though they concern what may be only a minor consideration in rain catch under the conditions described for the Koolau Mountains. It is perhaps just as reasonable to consider another approach to verify the validity of the rainfall as a function of distance for this situation. The equation determined for the rainfall (y) as a function of distance (x) over the interval between gages 1 and 10 is

$$y = ke^{ax}$$

where the constant (k) is the rainfall at gage 1 and the parameter (a) is equivalent to $(dy/dx)/y$, reflecting the incremental growth of rainfall with distance. This equation pertains not only

to the year 1957 covered in my paper but to the years 1958 and 1959, for which the data have now been evaluated. ('Rainfall and its distribution in the leeward Koolau Mountains, Hawaii,' *Pacific Science*, in press.) For each of the three years the value of (a) is statistically similar. This would appear to suggest that there is a regularity in the change in rainfall with distance that overshadows irregular changes expected from random wind conditions.

Beyond gage 10, where the slope of the terrain abruptly increases, the decrease in rainfall is unquestionably affected by wind and associated turbulence, but the linear distance is small (about 2000 ft) and it is just as likely that the gages reflect the true amount of rain falling on the ground as not. Because of the apparent irregularity of the rain catch in this area, an attempt was made to establish a distribution equation for the rainfall.

Evaporation Reduction by Natural Surface Films

JAMES W. DEARDORFF

*Department of Meteorology and Climatology
University of Washington, Seattle*

indication of the minimum amount by naturally occurring surface films reduce evaporation rate in waters of the San Juan of northwestern Washington was obtained through the use of a pair of evaporation

circular aluminum cake pans, each with height of 4 cm and an area of 134 cm², used. In the bottom of one was a hole, 1 cm in diameter, which could be corked from below. The procedure, which could be followed only when the sea was not too rough, was to sink this pan with the bottom uncorked and to raise the surface in such a manner as to leave as little as possible any surface film which may have been present. The pan was then raised toward the surface in a horizontal position so that its rim intercepted the water surface and also, presumably, any film which may have been present thereon. The pan was then lifted further and the water allowed to drain through the hole until a depth of only 1 mm remained, at which time the hole was corked from beneath. The pan was then placed within a frame inside of an automobile inner tube which floated on the surface. The frame supported the pan in such a manner that the pan was in contact with the natural water surface and its upper side was exposed to the air. The other pan was filled to approximately the same depth, but with water collected a short distance beneath the surface. This pan was also placed within a frame and floating inner tube similar to those already mentioned. The inner tubes were used in order to maintain the water surface within the pans as close as possible to that of the environmental water and yet prevent outside water from splashing into the

In addition, a sample of the sea water was collected and bottled at the time the two pans were used. This sample was taken for the purpose of determining the initial salinity of the water in both pans before evaporation occurred.

At the end of the test period the water within both pans was also collected so that the respective salinities, and hence evaporation rates, could subsequently be determined from the equation

$$E = (\rho V)_1(S_1 - S_0)/S_0 A \Delta t$$

where E is the evaporation rate, ρ the density of the water, V the volume of water, S the salinity, A the area exposed to evaporation, and Δt the time interval of the evaporation. The subscripts 0 and 1 refer to initial and final values, respectively. This equation follows from consideration of the conservation of total water mass and salt mass.

Since the ratio V/A was rather small for the pans used, large salinity differences were obtained over periods as short as 1 hour, the length of the evaporation period.

Several sets of measurements were taken in order to determine whether by this procedure the presence of naturally occurring surface films could be detected, and, if so, to determine whether such films are compressed enough to reduce the evaporation rate by a significant amount.

In all but one of the measurements the two pans were placed about 50 meters from the shore at the University of Washington Laboratories, Friday Harbor, Washington, during periods of light onshore winds. The pans were oriented cross-wind from each other and separated by a distance of 3 meters. In one case the test was made within East Sound, Orcas Island, at a distance of about $\frac{1}{2}$ km from shore, with an offshore wind of less than 1 m/sec. In this case the pans were not suspended within the inner tubes, but were placed atop a barrel on a raft and interchanged each 10 minutes to ensure nearly equal exposure to wind and solar radiation.

In all tests both pans were wiped clean just before each evaporation period in order to avoid contamination from more saline water or from a surface film which might have remained in

TABLE 1.

No.	Date	Location	Evaporation Rate, mm/day		Percentage Reduction ($E_1 - E_2$)/ E_1
			Pan with Sub- surface Water E_1	Pan with Sur- face Water E_2	
1	7/11/58	Friday Harbor	2.37	1.83	23
2	8/12/58	Friday Harbor	1.20	1.17	2
3	8/12/58	Friday Harbor	3.63	3.09	15
4	8/12/58	Friday Harbor	2.16	1.88	13
5	8/13/58	Friday Harbor	— .58	— .22	62
6	8/ 9/60	East Sound	4.19	3.31	21
Comparison Tests:					
7	8/12/58	Friday Harbor	2.30	2.10	9
8	8/13/58	Friday Harbor	1.10	1.17	— 6
9	8/13/58	Friday Harbor	.51	.50	2

the pan from a previous test. Care was also taken to avoid contamination or evaporation of the water samples during their collection, bottling, and salinity evaluation.

Results of six sets of measurements are given in Table 1. Tests 7 to 9 were taken with both pans filled with subsurface water for comparison purposes. Although relative differences between pan evaporation rates of up to nearly 10 per cent occurred in these three comparison tests (presumably a result of slight inequalities of exposure of the two pans), the larger differences which occurred in five of the six measurements, and in the expected sense, are considered to be significant in indicating the presence of surface films compressed sufficiently to reduce the evaporation rate. During measurements 1, 4, and 6 the presence of such a film in the vicinity of the pans was suspected from visual evidence, but during measurements 2, 3, and 5 its presence was not noted. Table 1 indicates that only during measurement 2, and possibly also during measurement 5, was a compressed film absent. However, the results of measurement 5 may be questioned because of the small condensation rates involved. Furthermore, in a controlled test taken later, when condensation rates were much larger and the water surface in one pan was covered with a film composed of docosanols and octadecanol alcohols, the two condensation rates were not significantly different.

During measurement 6 the presence of a compressed film was confirmed by swishing the water surface by hand until the outline of a fresh surface became visible. This hole in the film closed up within a few seconds after it had

been formed. During this measurement the surface film apparently occupied a large portion of East Sound, but just how large is not known because the wind speed may have been too low to cause noticeable ripples even in the absence of a film.

This method of detection of surface films is believed to yield percentage ratios of the evaporation reduction somewhat smaller than actually occurring for the following reasons. First, the evaporation rate of the film-covered areas of the natural water surface is probably larger relative to that of the film-covered pans than is indicated by the two pans because the natural water surface is rougher when the surface film is free. This difference in surface roughness cannot, of course, be duplicated within the floating evaporation pans. Second, the evaporation pans designed to capture the surface film presumably have a capture efficiency less than unity. Perhaps the surface film, when present, must climb the interior walls of the pan.

The rather large reduction of evaporation rate of the order of 20 per cent, apparently caused by naturally occurring surface films indicates that their frequency of occurrence and areal coverage may need to be estimated and, if possible, taken into account when evaluating the evaporation rate of inland waters by any method in which the surface is assumed to be uncondensed.

I should like to acknowledge the support of the National Science Foundation grant G3991 for the design and use of the floating evaporation pans.

(Received May 13, 1961.)

Discussion of Paper by R. Rangarajan,
'A New Approach to Peak Flow Estimation'

SIMEON M. BERMAN

Columbia University
New York, New York

theory of extreme values, as expounded in standard texts, has been developed under the assumption that the observations from which extreme values are generated are mutually independent [Gumbel, 1958]. This assumption is always necessary in order that the theory be correct; there may exist some dependency among observations. For example, it has been shown by Watson [1954] that the usual limiting distribution of the extreme values is valid even in the case of a sequence of 'm-dependent' random variables, that is, a sequence $\{x_n\}$ in which x_i is independent of x_j if $|i - j| > m$. In other words, a given random variable in the sequence depends on its 'neighbors' but not on those 'far' from it.

Watson's theorem and proof are stated in a rather esoteric manner and will now be recalled. He showed that if $\{x_n\}$ is a stationary sequence of m-dependent random variables and fixed $\xi > 0$ we define $C_n(\xi)$ by means of

$$\xi = nP\{x_1 > C_n(\xi)\}$$

for every even integer l ,

$$\frac{(-\xi)^q}{q!} \leq \lim_{n \rightarrow \infty} P\{\max_{i \leq n} x_i \leq C_n(\xi)\} \leq \sum_{q=0}^l \frac{(-\xi)^q}{q!}$$

since l is arbitrary,

$$\lim_{n \rightarrow \infty} P\{\max_{i \leq n} x_i \leq C_n(\xi)\} = e^{-\xi}$$

can be restated in more conventional language. It is tacitly assumed that the marginal distribution function $F(x)$ of $\{x_n\}$ is such that if $\{x_n\}$ were mutually independent, their common sum would have a limiting distribution $\Phi(x)$; precisely, there exist sequences $\{a_n\}$ and $\{b_n\}$ such that for all x ,

$$\lim_{n \rightarrow \infty} P\{\max_{i \leq n} x_i \leq a_n x + b_n\} = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = \Phi(x)$$

Since

$$F(a_n x + b_n) \rightarrow 1$$

it follows that

$$\begin{aligned} nP\{x_1 > a_n x + b_n\} &= n(1 - F(a_n x + b_n)) \\ &\sim -n \log F(a_n x + b_n) = -\log F^n(a_n x + b_n) \\ &\rightarrow -\log \Phi(x) \end{aligned}$$

Then the statement and proof of Watson's theorem can be illuminated by substituting

$$a_n x + b_n \text{ for } C_n(\xi)$$

and

$$-\log \Phi(x) \text{ for } \xi$$

It follows that the limiting distribution of the extreme value is $\Phi(x)$, the very same limiting distribution as in the case of independence; consequently the m-dependence does not affect the limiting distribution.

Although the analysis of floods has often been done by means of extreme value theory, there has been a question whether this procedure is theoretically sound; the daily discharges of a river are not, in fact, independent random quantities. Watson's result removes this objection. Although it is true that one day's discharge may be dependent on the discharges of the days that immediately precede or follow it, it is reasonable to assume that daily discharges far apart in time are indeed independent. The discharges, then, are an example of an m-dependent sequence of random variables; Watson's theorem applies, and the usual extreme value theory can be used.

The question of dependence was handled recently in the paper of Rangarajan [1960]; he showed that if the successive pairs of observations are mutually dependent, the usual extreme-value distribution still holds. It is clear, then, that Rangarajan's result is not new but is a special case of Watson's.

There are errors in the derivation which is presented by Rangarajan. He states that 'each x_i is correlated only with x_{i-1} ' and then writes this condition in the form

$$P\{x_i \leq X \mid x_{i-1}, x_{i-2}, \dots\} \\ = P\{x_i \leq X \mid x_{i-1}\} \quad (3)$$

This is an incorrect formulation. In reality condition 3 does not indicate that x_i depends only on x_{i-1} but instead indicates that x_i , $i = 1, 2, \dots$ forms a Markov chain. An example can be given in which all the x_i , $i = 1, 2, \dots$ are mutually dependent but where condition 3 does in fact hold; that is, the conditional distribution of x_i given x_i, x_{i-1}, \dots depends only on x_{i-1} . The example is the well-known stationary gaussian Markov process [Doob, 1953] where the $\{x_i\}$ have a joint normal distribution with

$$Ex_i = 0, Ex_i^2 = 1, Ex_i x_j = e^{-|i-j|}, \quad (i \neq j)$$

It is clear that these random variables are all mutually dependent; it is well known that they also satisfy condition 3 [Doob, 1953].

Since condition 3 does not characterize Rangarajan's stated conditions but instead characterizes the conditions of a Markov process, can it be said that he has proved his result for a Markov process? The answer is No, for there are inconsistencies between condition 3 and the supposed implications he draws from it. From

condition 3 he infers that

$$P\{x_n \leq X \mid x_{n-1} \leq X, \dots, x_1 \leq X\} \\ = P\{x_n \leq X \mid x_{n-1} \leq X\}$$

But (3') does not follow from (3). The form states that the conditional distribution of x_n given the values of x_1, x_2, \dots, x_{n-1} , depends on the value of x_{n-1} . On the other hand states that the conditional distribution of x_n if x_1, \dots, x_{n-1} have values not exceeding X depends only on the fact that x_{n-1} does not exceed X . It is known that (3') does not always hold in a Markov chain [Feller, 1957].

In summary, Rangarajan's result is not new but is a special case of a known general theorem and the derivation in his paper contains errors from the standpoint of the theory of probability.

I should like to thank Professor E. J. Gumbel for calling my attention to Rangarajan's paper.

REFERENCES

- Benson, M. A., Discussion of paper by R. F. Rangarajan, 'A new approach to peak flow estimation,' *J. Geophys. Research*, **65**, 2557, 1960.
 Doob, J. L., *Stochastic Processes*, John Wiley & Sons, New York, p. 218, 1953.
 Feller, William, *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons, New York, p. 379, example 27 (note), 1957.
 Gumbel, E. J., *Statistics of Extremes*, Columbia University Press, New York, 1958.
 Rangarajan, R., A new approach to peak estimation, *J. Geophys. Research*, **65**, 643-647, 1960.
 Watson, G. S., Extreme values in samples from m -dependent stationary stochastic processes, *Ann. Math. Stat.*, p. 798, 1954.

(Received May 3, 1961.)

A Note on the Accuracy of Drainage Densities Computed from Topographic Maps¹

WILLIAM J. SCHNEIDER

*U. S. Geological Survey
Washington, D. C.*

Criteria for the measurement of stream lengths have been set forth by Morisawa [1957]. They have been adopted in several geomorphic studies, including a recent study by Gray [1961].

Stream lengths are frequently expressed as a ratio of channel per square mile of drainage area. This ratio is termed the drainage density of a basin. In recent evaluation of drainage density as a hydrologic variable, data obtained by Morisawa [1957] were reanalyzed to determine the accuracy of drainage densities determined from topographic maps. Data consist of drainage densities for 18 basins computed by each of two methods: (1) by field measurement, (2) by reading drainage net on topographic sheet plotted on V-shaped crenulation in contours, and (3) by measuring streams as designated in blue

on the topographic maps. Of these 18 basins, eleven were computed for areas less than 0.5 square miles using maps at a scale of 1-24,000 and seven were computed for areas larger than 0.5 square miles using maps at a scale of 1-62,500.

Using paired *t* tests, Morisawa concluded that measurements of streams as indicated in blue on topographic sheets differ from actual field measurements of the streams, but that measurements made from drainage nets based on contour crenulations do not differ significantly from field measurements. The paired *t* test as used by Morisawa tested only the hypothesis that the mean of drainage densities computed by each method from topographic maps do not differ from the mean of the field measurements. No evaluation was made of other possible relationship or parameters.

Publication authorized by the Director, U. S. Geological Survey.

The first step in the re-evaluation of the data

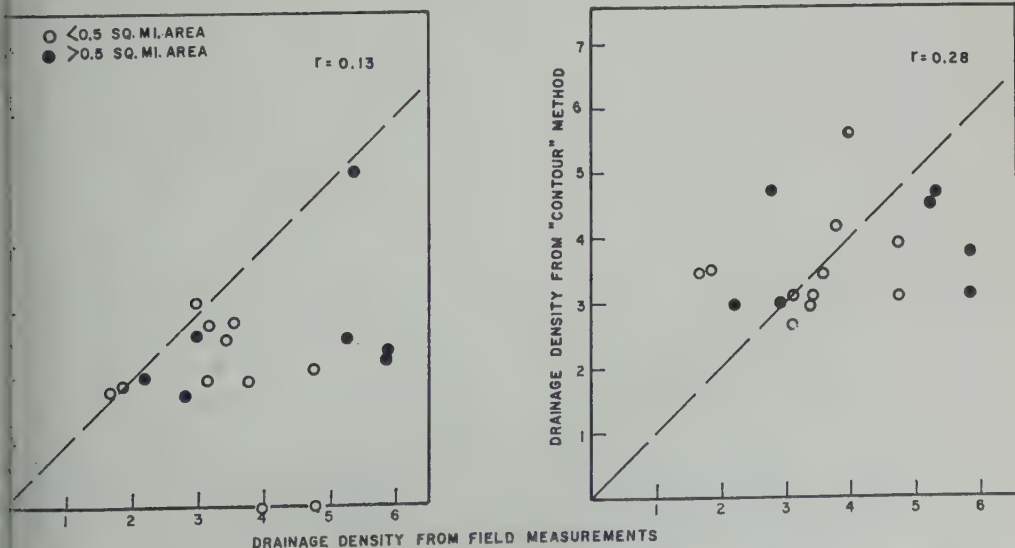


Fig. 1. Relation of drainage densities determined from topographic maps to those obtained by field measurements.

established that no significant differences were found at the 90 per cent level in either means or variances of the two groups of data which had been arbitrarily divided on the basis of size of area and map scale. The conclusion, therefore, is that neither map scale nor size of area affects the values of drainage density for the general area studied. Analyses of variance [Dixon and Massey, 1952, p. 145-155] indicate significant differences between stations but not between methods when the two methods of computing drainage densities are tested separately against the field measurements.

However, the ultimate test of the reliability of the values of drainage density determined from topographic maps is in the degree of correlation between these values and actual field measurements. Coefficients of correlation [Dixon and Massey, 1957, p. 200] were computed as 0.13 for the relationship between field measurements

and the 'blue line' method and 0.28 for the relationship between field measurements and the 'contour' method (Fig. 1). On the basis of the poor correlations alone, we must conclude that neither method of computing drainage density from topographic maps gives reliable estimates of the 'true' value as determined by field measurements.

REFERENCES

- Dixon, W. J., and Massey, F. J., *Introduction to Statistical Analysis*, 2nd ed., McGraw-Hill Book Co., New York, 1957.
- Gray, C. M., Interrelationships of watershed characteristics, *J. Geophys. Research*, 66, 12, 1223, 1961.
- Morisawa, M. E., Accuracy of determination of stream lengths from topographic maps, *Trans. Am. Geophys. Union*, 38, 86-88, 1957.

(Received March 23, 1961;
revised July 10, 1961.)

Reply to Letter by W. J. Schneider,

'A Note on the Accuracy of Drainage Densities Computed from
Topographic Maps'

M. MORISAWA

Geology Department, Montana State University, Missoula, Montana

I am glad that someone on the United States Geological Survey realizes that drainage density computed from topographic maps does not give a reliable estimate of the true value. This is what I tried to say in my paper, and have maintained. In fact, in all my work I do not use measurements from maps—even as marked by V-shaped contours—till heads are checked in the field.

However, it is not generally with the drainage density of each particular basin, but the

mean of an area, with which one works—at least, with which I and others have worked. This was the reason I tested the means rather than variation of each basin, as is reported in the present paper.

Again, let me commend the present author. I hope he can convince his fellow workers on the USGS that maps cannot be used for stream length or density computations.

(Received June 29, 1961.)

Discussion of Paper by W. E. Marlatt, A. V. Havens, N. A. Willits, and
G. D. Brill, 'A Comparison of Computed and Measured Soil Moisture
under Snap Beans'¹

R. M. HOLMES

*Agro-Meteorological Section, Canada Department of Agriculture
Ottawa, Canada*

It has been, and is, the search of investigators in the general area of soil-plant-moisture relationships to devise tools and techniques to obtain a space-integrated measure of soil moisture content and soil moisture losses. Meteorological methods have promise of accomplishing this. Marlatt, Havens, Willits, and Brill [1961] have made a significant contribution in this regard. Their work and that of their predecessors should stimulate research in this direction.

It may be of general interest that similar

¹ Contribution 151, Plant Research Institute, Research Branch, Canada Department of Agriculture.

work has been conducted in Canada [Holmes and Robertson, 1959, 1960; Holmes, 1961] based on early studies of Lehane and St. John [1953] and later by Kohler [1957] and Thwaithe and Mather [1954].

The work to which I refer primarily substantiates the results of Marlatt, Havens, Willits, and Brill [1961]. In fact, our soil moisture drying curves were a pattern of that shown by Marlatt in Figure 1 (dotted line) [Holmes and Robertson, 1959]. In a controlled environment experiment involving several soils, drying curves were found to be of the typical exponential

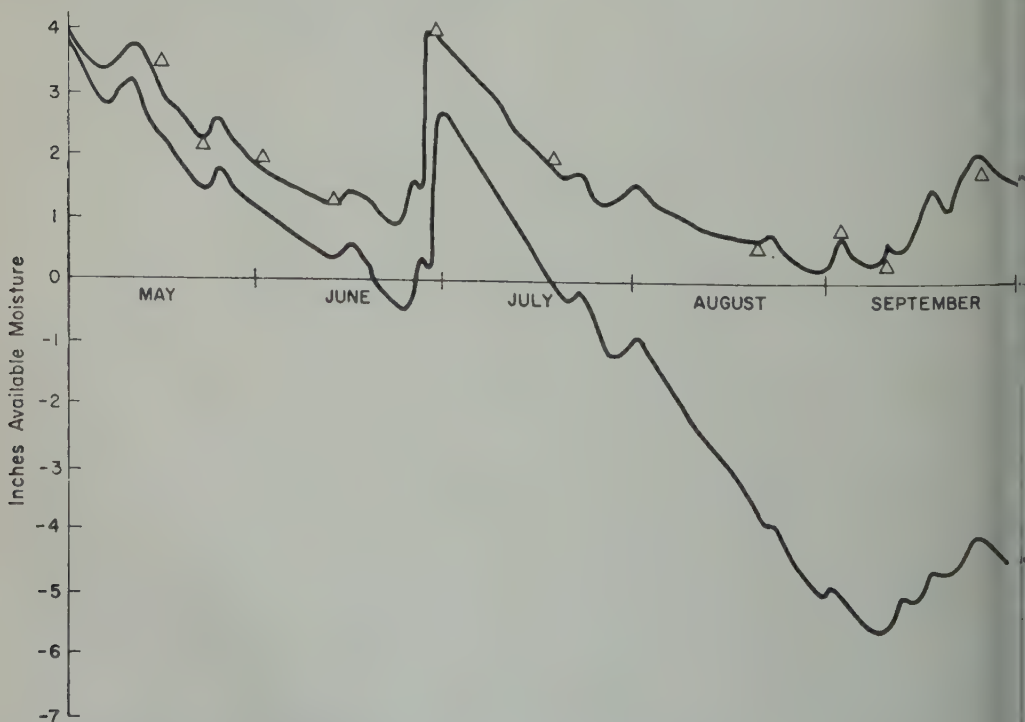


Fig. 1. Soil moisture loss calculated by modulated budget (AE), and by simple budget (PE); triangles indicate spot Coleman moisture block readings.

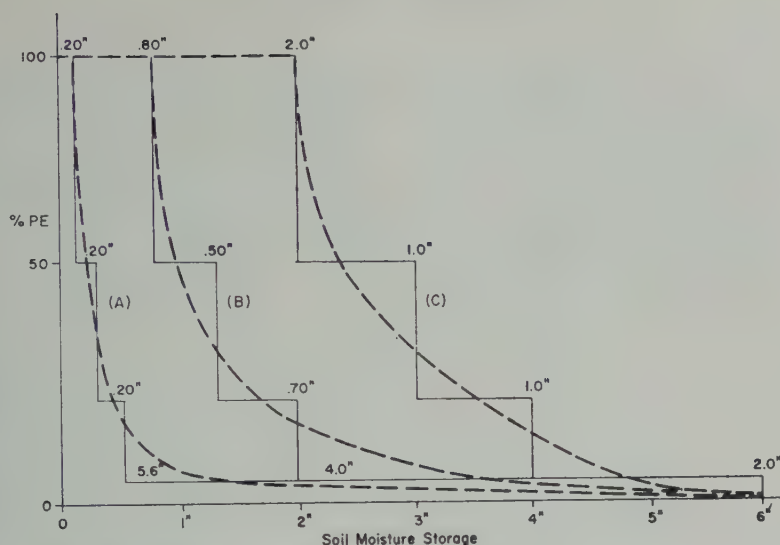


Fig. 2. IBM 650 program for adjustment of PE as soil dries and plant roots expand. (Curve A, Aug. 1 to May 31; curve B, June 1 to June 30; curve C, July 1, to July 31.)

pe, following an initial plateau where AE (actual evapotranspiration = potential transpiration). The length of time that PE was found to be strongly dependent texture, as suggested by Marlatt, and on rooting depth [Holmes and Robertson, Holmes, 1961].

order to make PE measurements useful estimating soil moisture content, it was necessary to multiply PE by a factor less than 1. A choice of factors made it possible to estimate soil moisture fairly accurately [Holmes and Robertson, 1959]. Figure 1 shows close agreement between Coleman moisture block readings and the modulated technique of estimating soil moisture, where PE was multiplied when $K < 1$. The PE curve is soil moisture estimated by a simple irrigation-type bud-

Other, we programmed the modulated program for an IBM 650 computer and computed soil moisture at Lethbridge, Alberta, for a 36-month period. The drying curves used in this simulation are shown in Figure 2 and were based on data from through field and laboratory observations. The dotted lines are the ideal curves and the stepped lines are a best fit designed for the computer. For example, from harvest (approximately August 1) to May 31, 0.20 inch of water was evapotranspired at 100 per cent of

the maximum rate and the following 0.20 inch at 50 per cent, etc. The computer shifted from one curve to the other at the appropriate time. Similarly, a simple irrigation-type budget was programmed, and computations of soil moisture were made for the same period.

From the above, PE , AE , and moisture deficit (inches/month) were obtained and related to wheat yields obtained from a continuous wheat rotation at Lethbridge, Alberta. It was found that soil moisture content computed by the modulated technique (curves in Fig. 2) was significantly related to wheat yields, while relationships with soil moisture computed with the simpler technique were not significant.

This work and that of Marlatt, Havens, Willits, and Brill [1961] and others point to the unique possibilities of meteorological methods in estimating soil moisture content. By proper adaptation to each situation in regard to soil, crop, and culture practice, it should be possible to make these methods extremely useful in many water-balance-control problems.

REFERENCES

- Holmes, R. M., and G. W. Robertson, A modulated soil moisture budget, *Monthly Weather Rev.*, 87(3), 101-106, 1959.
- Holmes, R. M., and G. W. Robertson, The relationship between actual and potential evapotranspiration with application to a problem in

- arid-zone agriculture, *Proc. Am. Soc. Agric. Eng.* no. 16-200, Columbus, Ohio, June 1960.
- Holmes, R. M., Estimation of soil moisture content using evaporation data. Symposium on Evaporation, University of Toronto, March 1961; subcommittee on Hydrology, Associate Committee on Geodesy and Geophysics—National Research Council of Canada, 1961.
- Kohler, M. A., Computation of evaporation and evapotranspiration from meteorological observations. U. S. Weather Bur. Am. Meteorol. Soc. meeting March 19-21, Chicago, Ill., 1957.
- Lehane, J. J., and W. J. Staple, Water retention and availability in soils related to drought assistance, *Can. J. Agric. Sci.*, **33**, 265-273, 1953.
- Marlatt, W. E., A. V. Havens, N. A. Willits, and G. D. Brill, A Comparison of computed and measured soil moisture under snap beans, *Geophys. Research*, **66**(3), 535-541, 1961.
- Thornthwaite, C. W., and J. R. Mather, *Water Balance*, Drexel Institute, Laboratory of Climatology, Centerton, N. J., 1954.

(Received May 31, 1961.)

Discussion of Paper by Jacob Bear, 'On the Tensor Form of Dispersion in Porous Media'

G. DE JOSSELIN DE JONG AND M. J. BOSSEN

*Civil Engineering Department, Technological University
Delft, The Netherlands*

Bear's approach nicely shows for the specific two-dimensional case that the coefficient of dispersion is a fourth-rank tensor. The question arose whether it is possible to construct a differential equation which describes this phenomenon. In fact, this may be done for a uniform velocity field by the following analysis.

Let a rectangular cartesian coordinate system x_1, x_2, x_3 be placed so that the y_1 coordinate has the direction of the uniform flow v . The lateral dispersion, i.e. in the direction of flow (y_1), with coefficient D_I is larger than the transverse dispersion D_{II} in the directions y_2, y_3 perpendicular to the flow. The rate of change of the concentration due to dispersion is therefore described by three terms: for the lateral dispersion, $|v| D_I \partial^2 C / \partial y_1^2$; for the transverse dispersion, $|v| D_{II} [\partial^2 C / \partial y_2^2 + \partial^2 C / \partial y_3^2]$; and for the convection, $-v \partial C / \partial y_1$. Combination of these terms gives

$$\frac{\partial C}{\partial t} = |v| \{ D_I \partial^2 C / \partial y_1^2 + D_{II} \partial^2 C / \partial y_2^2 + D_{II} \partial^2 C / \partial y_3^2 \} - v (\partial C / \partial y_1) \quad (1)$$

This relation can be written

$$\frac{\partial C}{\partial t} = |v| \{ (D_I - D_{II}) \partial^2 C / \partial y_1^2 + D_{II} \nabla^2 C \} - v (\partial C / \partial y_1) \quad (2)$$

Let the rectangular cartesian coordinates x_1, x_2, x_3 be rotated with respect to y_1, y_2, y_3 , so that $\alpha_1, \alpha_2, \alpha_3$ are the angles between y_1 and x_1, x_2, x_3 , respectively.

The chain rule for differentiation gives

$$\begin{aligned} \frac{\partial C}{\partial y_i} &= \sum (\partial x_i / \partial y_i) (\partial / \partial x_i) \\ &= \sum \cos \alpha_i (\partial / \partial x_i) \end{aligned}$$

where v_i is the component of v in the direction x_i ,

$\cos \alpha$ is v_i / v and

$$\partial / \partial y_1 = \sum (v_i / v) (\partial / \partial x_i)$$

Applied twice, this gives

$$\begin{aligned} \partial^2 C / \partial y_1^2 &= v^{-2} [v_1^2 C_{11} + v_2^2 C_{22} + v_3^2 C_{33} + 2v_1 v_2 C_{12} \\ &\quad + 2v_2 v_3 C_{23} + 2v_1 v_3 C_{13}] \\ &= v^{-2} \sum \sum [v_\alpha v_\beta C_{\alpha\beta}] \end{aligned}$$

where $C_{12} = \partial^2 C / \partial x_1 \partial x_2$. Since ∇^2 is invariant for rotations of rectangular cartesian coordinates and $\cos^2 \alpha_1 + \cos^2 \alpha_2 + \cos^2 \alpha_3 = \Sigma v_i^2 / v^2 = 1$, it is correct to write

$$\begin{aligned} \nabla^2 C &= v^{-2} [v_1^2 C_{11} + v_2^2 C_{11} + v_3^2 C_{11} \\ &\quad + v_1^2 C_{22} + v_2^2 C_{22} + v_3^2 C_{22} \\ &\quad + v_1^2 C_{33} + v_2^2 C_{33} + v_3^2 C_{33}] \end{aligned}$$

Therefore (2) becomes

$$\begin{aligned} \frac{\partial C}{\partial t} &= |v|^{-1} \{ D_I v_1^2 C_{11} + D_{II} v_2^2 C_{11} + D_{II} v_3^2 C_{11} \\ &\quad + D_{II} v_1^2 C_{22} + D_I v_2^2 C_{22} + D_{II} v_3^2 C_{22} \\ &\quad + D_{II} v_1^2 C_{33} + D_{II} v_2^2 C_{33} + D_I v_3^2 C_{33} \\ &\quad + 2(D_I - D_{II}) [v_1 v_2 C_{12} + v_2 v_3 C_{23} + v_3 v_1 C_{31}] \\ &\quad - v_1 C_1 - v_2 C_2 - v_3 C_3 \}. \end{aligned}$$

This can be abbreviated by the following expression:

$$\frac{\partial C}{\partial t} = \left\{ D_{ijkl} \frac{v_i v_j}{|v|} \frac{\partial^2 C}{\partial x_k \partial x_l} \right\} - \frac{\partial}{\partial x_i} (v_i C) \quad (3)$$

where, with $D^* = \frac{1}{2}(D_I - D_{II})$,

$D_{ijkl} =$

D_I	D_{II}	D_{II}	0	0	0	0	0	0	i	j
D_{II}	D_I	D_{II}	0	0	0	0	0	0	1	1
D_{II}	D_{II}	D_I	0	0	0	0	0	0	2	2
0	0	0	D^*	D^*	0	0	0	0	3	3
0	0	0	D^*	D^*	0	0	0	0	1	2
0	0	0	0	0	D^*	D^*	0	0	2	1
0	0	0	0	0	D^*	D^*	0	0	2	3
0	0	0	0	0	0	0	D^*	D^*	3	2
0	0	0	0	0	0	0	D^*	D^*	3	1
0	0	0	0	0	0	0	D^*	D^*	1	3

k	1	2	3	1	2	2	3	3	1
l	1	2	3	2	1	3	2	1	3

This expression (4) for the constant of dispersion is identical to the formula 53 [Bear, 1961, 1196].

REFERENCE

Bear, J., On the tensor form of dispersion in porous media, *J. Geophys. Research*, 66, 1185-1197, 1961.

(Received July 24, 1961.)

Moiré Patterns of the Membrane Analogy for Ground-Water Movement Applied to Multiple Fluid Flow

G. DE JOSSELIN DE JONG

*Civil Engineering Department, Technological University
Delft, the Netherlands*

Introduction. In the membrane analogy for flow through porous media, contour lines of a deflected membrane represent either streamlines or equipotential lines. The membrane was used by Prandtl [1903] to solve torsion problems. Lichtenberg [1952] gave a description of the application of the analogy to solve the flow patterns resulting from systems of sources and sinks. Multiple fluid flow through porous media may be treated by considering a suitable distribution of sources and sinks [de Josselin de Jong, 1960] and therefore the membrane analogy is also applicable to this problem.

The object of this letter is to point out that the problems may be solved by a moiré method, and this technique provides a convenient procedure for establishing contour lines of a deflected membrane. The procedure is a slight modification of the method initiated by Lichtenberg [1955], who superimposed two photographic exposures of a loaded and an unloaded model. Lichtenberg used the reflection of the grid as seen in the mirrored model surface to obtain a moiré pattern which is a measure of surface deflection. The modified procedure differs in that the grid pattern is projected on the nonreflecting membrane and the moiré pattern obtained by the superimposed images is a measure of the normal displacement. Otto [1954] mentioned this kind of experiment in connection with the study of deformed roofs.

Membrane-moiré test setup. The membrane is a thin rubber sheet placed vertically to eliminate the effect of its own weight and stretched uniformly by a constant force S . Sinks or sources are simulated by concentrated point loads P normal to the plane of the membrane as shown in Figure 1.

A horizontal beam of parallel light rays, at an angle of incidence α with the normal to the deflected membrane, is projected through a grid and is thrown a shadow image on the membrane. The grid consists of a ruling of parallel black

lines on a glass plate, oriented in such a way that the lines are vertical.

When the membrane is given a small displacement, w , normal to its plane, the shadow image of the grid is displaced. If the axis of the camera is placed normally to the unloaded membrane the apparent normal displacement, u , of the projected grid lines is

$$u = w \tan \alpha$$

With b the pitch of the shadow lines in the unloaded condition, a displacement of n spacings corresponds to a deflection w_n equal to

$$w_n = nb \cot \alpha$$

If the spacing of the grid lines is such that the pitch is twice the width of the lines, photographic superposition will produce complete exposure in those regions where w' is given by

$$w' = (n + \frac{1}{2})b \cot \alpha \quad (1)$$

therefore these regions are black in the negative and show up as white bands in the positive prints.

These bands indicate the desired contour lines for the deflected membrane. An example of the results which can be obtained by this method is given in Figure 2. This figure shows the streamlines for a confined aquifer filled with two fluids of different specific weight at the moment the fluid motion starts from an abrupt vertical interface. Since this is the same problem as treated previously by the author (1960), the upper half of Figure 2 can be compared with the full lines of Figure 15 of the article quoted, wherein the results of the electric resistance model are represented.

Computation of analogous quantities. The dimensions of the test setup were determined as follows. The relation between the deflection w and the normal load per unit area p of a membrane uniformly stretched by a unit force S is

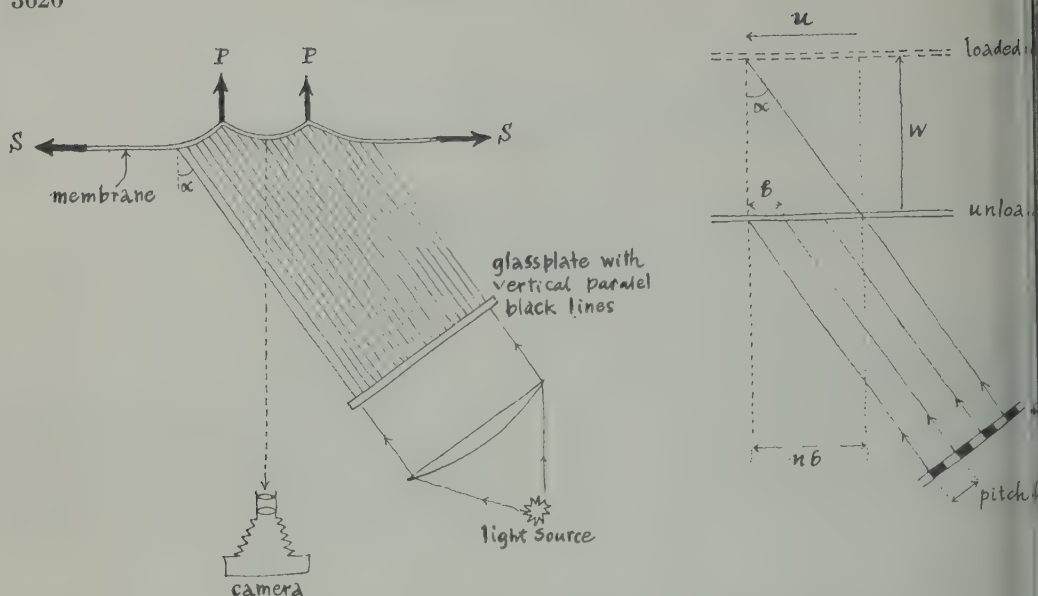


Fig. 1. Top view of test arrangement and detail of displaced membrane.

given by the well-known equation

$$\nabla^2 w = -p/S \quad (2)$$

[e.g. see *Timoshenko and Goodier*, p. 269, 1951.]

This equation is analogous to the basic equation for multiple fluid flow given by the author in the reference quoted. Equating (10) to (11) from this reference gives

$$\nabla^2 \Psi = \omega \quad (3)$$

where Ψ is the specific discharge stream function and ω is the vorticity. Therefore, a vortex of strength ωdA acting in the region dA can be represented by a distribution $(-p/S)dA$ over the region dA of the membrane. The force distribution p may be approximated by a concentrated force, $P = -pdA$, applied at the centroid of the region dA . The vortex strength for an abrupt interface is given by equation 22 of the article quoted in corrected form:

$$\omega dA = \omega ds dn = (k/\mu)(\gamma_2 - \gamma_1) dy_0$$

From (2) and (3) the relationship between analogous quantities is therefore established:

$$\frac{w}{\Psi} = \frac{-p/S}{\omega} = \frac{-p dA/S}{\omega dA} = \frac{P/S}{(k/\mu)(\gamma_2 - \gamma_1) dy_0}$$

For the electric resistance model the streamline interval was arbitrarily selected as (eq. (1))

$$\Delta \Psi = (k/\mu)(\gamma_2 - \gamma_1)(c/20)$$

To permit comparison of test results the contour interval for the moiré test was therefore chosen to be

$$\Delta w = \Delta \Psi \frac{P/S}{(k/\mu)(\gamma_2 - \gamma_1) dy_0} = \frac{Pc}{20S dy_0}$$

Since nine point loads were used, for a half aquifer height c in the membrane, dy_0 equals $c/9$, and

$$\Delta w = 9P/20S$$

In order that Δw correspond to one interval n is 1 in (1), and Δw of (4) should be equal to $b \cot \alpha$. Thus it follows that

$$P = (20/9)Sb \cot \alpha$$

Using a stretching force $S = 200$ g/cm, projected grid width $b = 0.22$ cm and an angle of incidence $\alpha = 45^\circ$, point loads of magnitude $P = 98$ g were required. Since the theory is only valid for small values of w a correction had to be made for regions of deviation.

Figure 3 gives the pattern for the instantaneous streamlines corresponding to a subsequent position of the interface occurring at a time $t = 1/2 \epsilon \tau$.



Fig. 2. Moiré pattern of streamlines in two fluids of different weight filling a confined aquifer and separated by a vertical interface.

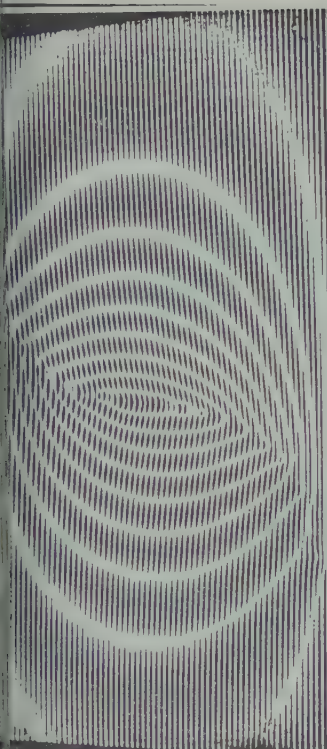


Fig. 3. Moiré pattern of streamlines for the displaced interface.

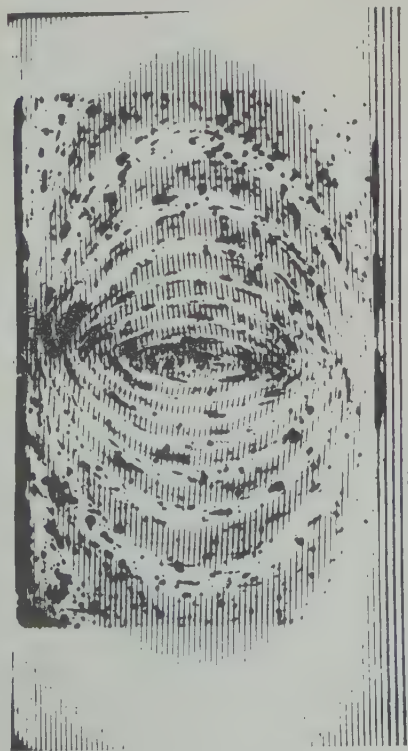


Fig. 4. Comparison of moiré pattern and parallel plate test for the vertical interface.

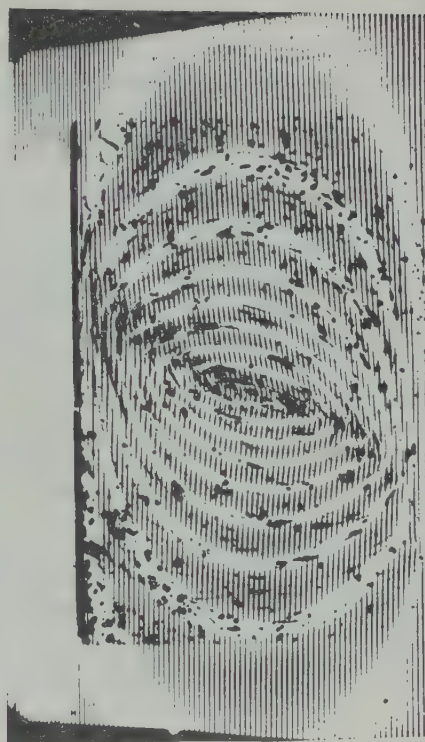


Fig. 5. Comparison of moiré pattern and parallel plate test for the displaced interface.

$k(\gamma_2 - \gamma_1)$ sec, after the initial vertical position of the interface given in Figure 2. The position of the interface at a subsequent instant may be approximated from the streamline pattern for the previous instant, assuming steady-state flow during the time interval.

In the membrane analogy it is then necessary to shift the point loads to the new position of the interface. For a complete investigation of the movement of the fluids, a succession of photographs representing the instantaneous motion in the entire field at each of the time intervals selected is required.

The membrane moiré results of Figures 2 and 3 are comparable to the parallel plate test results given in Figures 13 and 14 of the article quoted. Their agreement is demonstrated by the superposition shown in Figures 4 and 5.

Equipotential lines may also be obtained by the moiré method. Point load moment couples are then applied to the membrane to represent the doublets in the analogy.

The principal advantage of the membrane

moiré method is the convenient way of obtaining a photograph of streamlines and equipotentials. This procedure eliminates the need for point-to-point plotting of contour lines, which is required with the electrical resistance analogy.

REFERENCES

- Hansen, V. E., Complicated well problems solved by membrane analogy, *Trans. Am. Geophys. Union*, 33, 912-916, 1952.
- de Josselin de Jong, G., Singularity distribution for the analysis of multiple-fluid flow through porous media, *J. Geophys. Research*, 65(1), 3739-3758, 1960.
- Ligtenberg, F. K., The moiré method—A new experimental method for the determination of moments in small slab models, *Proc. Soc. Exptl. Stress Anal.*, 12(2), 1955.
- Otto, F., *Das Hängende Dach*, Ullstein A. G., Berlin, p. 22, 1954.
- Prandtl, L., Zur Torsion von prismatischen Stäben, *Physik. Z.*, 4(26b), 1903.
- Timoshenko, S., and J. N. Goodier, *Theory of Elasticity*, 2nd ed., p. 269, 1951.
- Todd, D. K., *Groundwater Hydrology*, John Wiley & Sons, New York, (chap. 14, p. 316: 'Membrane Models'), 1959.

Corrigendum

Professor G. de Josselin de Jong has called attention to the following errors in his paper 'Potentiality Distribution for the Analysis of Multiple Fluid Flow through Porous Media,' published in the November 1960 issue of the *Journal*. On page 3746, formula 22 should be:

$$\omega \, ds \, dn = (k/\mu)(\gamma_2 - \gamma_1) \, dy_0$$

On page 3756, the sentence beginning on line 20 of column 1 should be: 'A potential divider with 1000 subdivisions was supplied with a potential difference of $\frac{1}{24}E^*$.'

Information for Contributors to the *Journal of Geophysical Research*

Manuscripts. Send manuscripts to J. A. Peoples, Jr., Department of Geology, University of Kansas, Lawrence, Kansas. Manuscripts, including proof copies of figures, should be submitted in triplicate for expedite review and publication. Manuscripts should be in English, typewritten on heavy paper on one side of page only, double spaced (including abstracts and references), with generous margins.

Ample space should be allowed for mathematical expressions, which should be typed or very plainly written by hand. Particular attention should be given to the legibility of subscripts and superscripts and to differentiation between capital and lower-case letters. Unusual symbols and cumbersome notation should be avoided. Fractional exponents should be written in preference to root signs, and the solidus should be used for fractions wherever its use will save vertical space.

Authors are urged to have their papers critically reviewed by their associates for scientific validity, manner of presentation, and use of English before submitting them for publication.

Abstracts. An abstract must accompany each manuscript. It should be a concise but comprehensive condensation of the essential parts of the paper, suitable for separate publication and adequate for preparation of general indexes to geophysical literature.

References and footnotes. References should be indicated in the text by the insertion in brackets of the author's name and the year of publication, thus: [Faust, 1958]. If the author's name is part of the text only the year is bracketed. If there are two or more references citing different papers published in the same year by the same author, they may be distinguished by the letters *a*, *b*, *c* after the year.

At the end of the paper references should be listed alphabetically by the author's names in the form of the references given below. (See *List of Periodicals Abstracted by Chemical Abstracts* for abbreviations of titles of journals, or write titles in full.)

Footnotes to the text should be avoided; parenthetical sentences should preferably be added to the text. If footnotes must be given they should be numbered consecutively throughout the paper.

Tables and figures. Material suited to tabular form should be arranged as a table and may be typewritten on a separate sheet. Tables must be numbered according to their sequence in the text, and each table should have a title. Column headings should be short and self-explanatory; more complete explanations may be given in footnotes to the table.

Figures should be prepared with the column width of this Journal in mind (a scale of 2 to 4 times that of the published figure is usually adequate). They must be drawn in India ink on white paper or tracing cloth. Coordinate paper should be avoided; if used, however, it must be blue-lined, and coordinate lines that are to show must be inked.

Titles of figures must be typewritten consecutively on a separate sheet; they are *not* to be lettered on the figure. Necessary lettering within the figure must be executed to meet competent drafting standards and should be large enough to remain legible after reduction. The ideal letter size in the reduced figure is 1/16 inch, but 1 mm is acceptable. Unnecessary blank space within the figure should be avoided.

Photographs are acceptable only if they have good intensity and contrast. They should be unmounted glossy prints.

Figure number and author's name should be written lightly in pencil on each figure. 'Top' of each figure should be indicated.

Acknowledgments. Acknowledgments should be made only for significant contributions by the author or professional associates. A brief closing statement will usually suffice.

REFERENCES

- American Chemical Society, *List of Periodicals Abstracted by Chemical Abstracts*, Chemical Abstracts Service, Ohio State Univ., Columbus, 314 pp., 1956.
- American Institute of Physics, *Style Manual*, 2nd ed., American Institute of Physics, New York, 42 pp., 1959.
- Faust, L. Y., The preparation of a paper, *Geophysics*, 23, 944-952, 1958.
- Landes, K. K., A scrutiny of the abstract, *Geophysics*, 17, 645, 1952. Reprinted in *Geophysics*, 23, 942-943, 1958.
- Skillen, M. E., R. M. Gay, and others, *Words into Type*, Appleton-Century-Crofts, New York, 585 pp., 1948.
- U. S. Geological Survey, *Suggestions to Authors of the Reports of the United States Geological Survey*, 5th ed., U. S. Govt. Printing Office, Washington, 255 pp., 1958.
- William Byrd Press, *Mathematics in Type*, William Byrd Press, Richmond, 58 pp., 1954.



CORPORATION AND SUPPORTING MEMBERSHIPS

The American Geophysical Union is a non-profit scientific organization established by the National Research Council. Its Council is the United States National Committee of the International Union of Geodesy and Geophysics; official adherence by the United States is through the National Academy of Sciences-National Research Council.

Excerpts from the Statutes:

Article 3. Membership—The membership of the American Geophysical Union shall be as follows:

- (e) *Corporation Members*—Corporations and other organizations interested in geophysics elected by the Council of the Union. The designated representative of each such organization shall enjoy the privileges of a Member.
- (g) *Supporting Members*—Corporations, other organizations, and individuals interested in geophysics and desirous of supporting the Union may become members under the following classifications upon election by the Council.

(Continued on next page)

Cut along this line

American Geophysical Union

PROPOSAL FOR _____ MEMBERSHIP

Council, American Geophysical Union
1515 Massachusetts Avenue, N.W., Washington 5, D. C.

I, _____

in order to give an indication of our interest in the aims and activities of the American Geophysical Union, and to assist in maintaining and extending its program of publication and research work in the development of the geophysical sciences, the undersigned applies for _____ Membership in the AGU and, until further notice, agrees to pay dues at the rate established for this classification of membership, in accordance with the information set forth above and on the following page.

My name or Organization _____

(Signature)

Title

(over)

(Continued from previous page)

of the Union . . . : Contributing Members, . . . Sustaining Members, . . . Benefactors. . . . Each Supporting Member and the designated representative shall enjoy the privileges of a Member.

Corporation Members shall pay dues of not less than \$100 for each calendar year. Dues for Supporting Members shall be as follows:

Contributing Members	\$500
Sustaining Members	\$1000
Benefactors	\$5000

Lists of Corporation Members and Supporting Members will be published in each issue of the *Transactions*, and will be included in the Membership Directory as distributed units.

By Laws provide that one copy of each issue of the *Transactions*, *Journal of Geophysical Research*, any published *List of Members and Officers*, and any other publication which may be approved for free distribution to the membership shall be sent to each Corporation and Supporting Member. Each organization in good standing may purchase any available publication of the Union at the established member discount.

AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Ave., N.W.

Washington 5, D. C.

Cut along this line

Address _____

City _____ State _____

General fields of activity _____

The following person is designated as our representative in this membership _____

_____ Title _____

Number of units of membership desired (this will be taken as one unless otherwise indicated) _____

Place _____

Date _____

AMERICAN
GEOPHYSICAL
UNION

UNSELFISH
COOPERATION
IN RESEARCH

AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W., Washington 5, D. C.

Established by the National Research Council in 1919 for the development of the science of geophysics through scientific publication and the advancement of professional ideals.

QUALIFICATIONS FOR MEMBERSHIP

Membership of the AGU shall consist of Members, Associate Members, Student Members, Honorary Members, and Supporting Members.

Those eligible as candidates for election to the grade of MEMBER shall be:

MEMBER (a) Persons who have made an active contribution to geophysical research through observation, publication, teaching, or administration. Definite evidence should be presented to the Membership Committee. "Publication" may include books, articles, unpublished manuscripts, inventions, or development of geophysical instruments.

(b) Persons who have made active practical application of geophysical research. It should be shown that the nominee's work has not been purely routine, but that it has tended to create new knowledge of, or to broaden or strengthen the application of, geophysical research. In general, the minimum qualifications for membership will be not less than three years of professional experience in some phase of geophysics.

(Continued on next page)

Cut along this line

APPLICATION FOR MEMBERSHIP

Refer to qualifications on reverse side and designate below type of membership desired:

Member (\$10) ☐

Associate (\$10) ☐

Student (\$4.50) ☐
(1961)

Application forms for Corporation and Supporting Membership are available upon request.

Surname First Name Middle Name

Preferred mailing address for publications

Permanent address

Place Month Day Year of Birth 4. Country of citizenship/naturalization

Nature of work and title and/or military rank; name and address of organization with which you are associated.

Check section or sections with which affiliation is desired.

☐ Geodesy

☐ Seismology

☐ Meteorology

☐ Geomagnetism and Aeronomy

☐ Oceanography

☐ Volcanology, Geochemistry, and Petrology

☐ Hydrology

☐ Tectonophysics

EXPERIENCE (List below, use added sheets as necessary)

Dates: From To Name and address of organization Title, duties, nature of work

EDUCATION (List Below, use added sheets as necessary)

Dates: From To School Address Major Subject Degree, if any; year
(over)

(Continued from previous page)

Those eligible as candidates for election to the grade of ASSOCIATE MEMBER shall be:

ASSOCIATE MEMBER Persons who have an active interest in physical processes of the Earth or technical assistance in the application of geophysics. In general, the minimum qualifications for associate membership will be acceptable training or experience in some field of geophysics or allied science.

CORPORATION AND SUPPORTING MEMBER Corporations and other interested organizations shall be eligible as candidates for election to CORPORATION or SUPPORTING MEMBERSHIP. They shall have the privilege of designating a representative who has the rights and privileges of Members (use special form).

STUDENT MEMBER Those eligible as candidates for election to the grade of STUDENT MEMBER shall be persons who are graduate or undergraduate students in residence at least half-time and who are specializing in the geophysical sciences. Teaching or research assistants enrolled in more than half of a full-time academic program may also be eligible for Student Membership. Student Members shall have all the privileges of Members except that they shall not vote or hold office.



Cut along this line

*9. References: Please list below names and addresses of two or three references; include members of the AGU or others who know you well.

*10. Titles of technical contributions or publications, particularly those in the geophysical sciences, where published.

*11. Brief statement of any special interests or qualifications in the geophysical sciences.

Date _____

Written Signature _____

12. (STUDENT MEMBERS ONLY) The person whose signature appears above is known to _____ and is a student majoring in _____ (subject) at _____

(Name of college or university) expected to graduate in _____ (year) with the degree of _____

☐ He is a full-time student, or ☐ a teaching or research assistant enrolled in more than half full-time academic program.

(Signature of faculty sponsor) _____

☐ Check here if faculty sponsor is a member of _____ and willing to act as a regular sponsor for associate membership as well.

(Typed or printed name of sponsor) _____

(Title) _____

* Applicants for student membership may omit Questions 9, 10, and 11, but must fill in Question 12. Please return form with check or money order payable to American Geophysical Union, 1515 Massachusetts Ave., N.W., Washington 5, D. C.

Contents

(Continued from back cover)

	PAGE
Generalized Empirical Analysis of Cratering.....	Charles E. Violet 3461
Experimental Determination of Earthquake Fault Length and Rupture Velocity.....	Frank Press, Ari Ben-Menahem, and M. Nafi Toksöz 3471
Optimization from Explosions in Soil.....	C. Kisslinger, E. J. Mateker, Jr., and T. V. McEvilly 3487
Results of the 1960 Expedition to Krakatau.....	Robert W. Decker and Djajadi Hadikusumo 3497
Seismological Phase Velocities for a Lunar Seismic Experiment.....	Bruce A. Bolt 3513
Measurement of Cosmic-Ray Intensity in the Meteorites.....	J. R. Arnold, M. Honda, and D. Lal 3519
Gamma-Ray-Induced Radioactivity in Terrestrial Materials.....	Rama and Masatake Honda 3533
Reactive Species Produced by Cosmic Rays in Bruderheim and Other Stone Meteorites.....	M. Honda, S. Umemoto, and J. R. Arnold 3541
Neon, Argon 37, and Argon 39 in the Bruderheim Meteorite.....	E. L. Fireman and J. DeFelice 3547
Radioactivity of the Bruderheim Chondrite.....	M. W. Rowe and M. A. Van Dilla 3553
Petrography and Chemical Composition of the Bruderheim Meteorite.....	Michael Duke, Donald Maynes, and Harrison Brown 3557
Abstracts of the Papers Presented at a Symposium on Ground Water, Portland, Oregon, November 16-17, 1960, Sponsored by the Section of Hydrology, American Geophysical Union.....	J. Virginia Lincoln 3565
Geomagnetism and Solar Data.....	J. Virginia Lincoln 3571
Letters to the Editor:	
On the Bruderheim Meteorite.....	H. Baadsgaard, F. A. Campbell, R. E. Folinsbee, and G. L. Cumming 3574
Neon in the Bruderheim Meteorite.....	W. B. Clarke and H. G. Thode 3578
On the Sulfur Isotope Abundances in Abee and Bruderheim Meteorites.....	M. Shima and H. G. Thode 3580
On the Trace Element Abundances in the Bruderheim Meteorite.....	W. D. Ehmann, A. Amiruddin, P. R. Rushbrook, and M. E. Hurst 3581
On the Origin of Excess Xe^{129} in Stone Meteorites.....	P. M. Jeffery and J. H. Reynolds 3582
On the Cosmic-Ray-Produced V^{50} and K^{40} in the Iron Meteorite Aroos.....	Heinz Stauffer and Masatake Honda 3584
On the Stability of the Interface between the Solar Wind and the Geomagnetic Field.....	A. J. Dessler 3587
On the Propagation of Hydromagnetic Propagation and Geomagnetic Field Stability.....	P. J. Coleman, Jr., and C. P. Sonett 3591
On the Lifetime of Radiation Belt Protons with Energies between 1 Kev and 1 Mev.....	Harold Liemohn 3593
On the Discussion of the Letter by R. A. Hanel, 'Determination of Cloud Altitude from a Satellite'.....	G. Yamamoto and D. Q. Wark 3596
On the Evaluation of the Special World Interval Program during the IGY.....	Martin E. Nason 3597
On the Geomagnetic Field Micropulsations and Electron Bremsstrahlung.....	Wallace H. Campbell 3599
On the Short-Term Phase Perturbations Observed at 18 Kc/s.....	C. F. Sechrist and K. D. Felperin 3601
On the Comments on K. Vozoff's Paper 'Calibration of Pulsation Detector Coils'.....	James R. Wait 3603
On the Note on the Autocorrelation Coefficient of K_p and Its Relation to M Regions.....	W. R. Piggott 3604
On the Author's Comments on the Previous Discussion.....	Fred Ward 3605
On the Earth's Free Oscillations Observed on Earth Tide Instruments at Tiefenort, East Germany.....	W. Buchheim and S. W. Smith 3608
On the Discussion of Paper by John F. Mink, 'Distribution Pattern of Rainfall in the Leeward Koolau Mountains, Oahu, Hawaii'.....	G. Reynolds 3611
On the Author's Reply to the Preceding Discussion.....	John F. Mink 3612
On the Evaporation Reduction by Natural Surface Films.....	James W. Deardorff 3613
On the Discussion of Paper by R. Rangarajan, 'A New Approach to Peak Flow Estimation'.....	Simeon M. Berman 3615
On the Note on the Accuracy of Drainage Densities Computed from Topographic Maps.....	William J. Schneider 3617
On the Reply to Letter by W. J. Schneider, 'A Note on the Accuracy of Drainage Densities Computed from Topographic Maps'.....	M. Morisawa 3619
On the Discussion of Paper by W. E. Marlatt, A. V. Havens, N. A. Willits, and G. D. Brill, 'A Comparison of Computed and Measured Soil Moisture under Snap Beans'.....	R. M. Holmes 3620
On the Discussion of Paper by Jacob Bear, 'On the Tensor Form of Dispersion in Porous Media'.....	G. de Josselin de Jong and M. J. Bossen 3623
On the Soiré Patterns of the Membrane Analogy for Ground-Water Movement Applied to Multiple Fluid Flow.....	G. de Josselin de Jong 3625
On the Appendix.....	G. de Josselin de Jong 3629

Contents

- Variations of the Cosmic Radiation in November 1960. *J. A. Lockwood and M. A. Shea*
Balloon Measurement of Solar Cosmic Rays at Fort Churchill, Canada, during July 1959 *James A. Earl*
- The Calculation of the Electron Density in the Ionosphere from Elevation-Angle Measurements on Artificial Satellites *J. E. Titheridge*
An HF Radar Search for Possible Effects of Earth Satellites upon the Upper Atmosphere *T. A. Croft and O. G. Villard, Jr.*
- Expected Influence of a Localized Change of Ionosphere Height on VLF Propagation *James R. Wait*
- The Interpretation and Synthesis of Certain Spread-F Configurations Appearing on Equatorial Ionograms *Wynne Calvert and Robert Cohen*
A Waveguide Interpretation of "Temperature-Latitude Spread F" on Equatorial Ionograms *M. L. V. Pitteway and Robert Cohen*
- On the Index of Refraction in the Ionosphere. *Otto Theimer and Leonard S. Taylor*
Solar-Flare Effects on 2.5 and 5.0 Mc/s Atmospheric Radio Noise. *John R. Herman*
Infrared and Reflected Solar Radiation Measurements from the Tiros II Meteorological Satellite *W. R. Bandeen, R. A. Hanel, John Licht, R. A. Stampfl, and W. G. Stroud*
- A Statistical Study of Lower Atmospheric-Ionospheric Coupling *Norman J. Macdonald and Robert W. Knecht*
Upper-Atmosphere Structure Measurement Made with the Pitot-Static Tube *J. E. Ainsworth, D. F. Fox, and H. E. LaGow*
- Total Atmospheric Ozone and Geomagnetic Activity. *S. J. Ahmed and A. Halim*
Airborne Filters for the Measurement of Atmospheric Space Charge *C. B. Moore, B. Vonnegut, and F. J. Mallahan*
- Natural Radioactivity in the Atmosphere. *Rama and M. Honda*
Ground-Conductivity Determinations at Low Radio Frequencies by an Analysis of the Sferic Signatures of Thunderstorms. *J. R. Johler and C. M. Lilley*
- S_0 and Ocean *Tsuneji Rikitake*
Backscattering of 3.21-Centimeter Radiation by Water Bubbles *Louis J. Battan and Benjamin M. Herman*
- On the Vertical Circulation of the Mediterranean Sea. *Georg Wüst*
General Theory of Dispersion in Porous Media. *A. E. Scheidegger*
Statistical Geometry of Porous Media. *H. D. Fara and A. E. Scheidegger*
- Effect of Interstation Correlation on Regression Analysis *Nicholas C. Matalas and Manuel A. Benson*
- The Frequency Distribution of Near Extremes. *L. E. Borgman*
Reduction of Transpiration. *W. J. Roberts*
- Average Antecedent Temperatures as a Factor in Predicting Runoff from Storm Rainfall *Charles D. Hopkins, Jr., and Dale O. Hackett*
- An Improved Statistical Model for Evaluating Parameters Affecting Water Yields of River Basins *B. Harris, A. L. Sharp, A. E. Gibbs, and W. J. Owen*
- Changes in the Levels of Lakes Michigan and Huron. *Ivan W. Brunk*
Some Aspects of the Application of the Theory of Sediment Transportation to Engineering Problems *J. Bogárdi*
- Sediment Transport of Glacier-Fed Streams in Alaska. *Whitney M. Borland*
Variation of Soil Erodibility with Geology, Geographic Zone, Elevation, and Vegetation Type in Northern California Wildlands. *J. E. André and H. W. Anderson*
- Evaluating Wells and Aquifers by Analytical Methods. *William C. Walton and William H. Walker*
- Cratering Symposium:
- New Evidence for the Impact Origin of the Ries Basin, Bavaria, Germany *E. M. Shoemaker and E. C. T. Chao*
 Analysis of the Formation of Meteor Crater, Arizona: A Preliminary Report. *R. L. Bjork*
 High-Explosive Craters in Desert Alluvium, Tuff, and Basalt *B. F. Murphey and L. J. Vortman*
 Throwout Calculations for Explosion Craters. *Wilmot N. Hess and M. D. Nordyke*
 Pacific Craters and Scaling Laws. *R. B. Vaile, Jr.*
 Nuclear Craters and Preliminary Theory of the Mechanics of Explosive Crater Formation *M. D. Nordyke*

(Continued inside back cover)